

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 07:13:39 ; Search time 1836.07 Seconds  
(without alignments)  
12110.869 Million cell updates/sec

Title: US-09-245-198a-3  
Perfect score: 1373  
Sequence: 1 atgcattgttgaattga.....gacaaatgttgaataatg 1373

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

EST.\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hlc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hlc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vtl:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	765.8	55.8	834	13	BI766766 603056866
2	753.2	54.9	948	14	BO707185 AGENCOURT
3	747.2	54.4	963	14	BO671259 AGENCOURT
4	743.4	54.1	777	13	BI819200 603034614
5	725.4	52.8	1071	14	BM921213 AGENCOURT
6	701.4	51.1	731	13	BI871711 603395825

7	688	50.1	828	13	BI596681 603243254
8	677.4	49.3	728	13	BI870393 603395641
9	584.4	42.6	607	13	BI966060 1E72904.X
10	581.4	42.3	1027	17	AF163779 AF163779
11	562	40.9	940	14	BO884231 AGENCOURT
12	534.6	38.9	568	14	BM971606 BI-CEC1
13	531.6	38.7	910	12	BG110063 602279667
14	516.8	37.6	531	13	BI824443 603038693
15	493.6	36.0	918	12	BE577781 602092080
16	458.2	33.4	1033	11	AK020909 Mus muscu
17	454.4	33.1	561	10	AW763237 ur70d09.y
18	444	32.3	951	14	BO674188 AGENCOURT
19	442.8	32.3	587	12	BG668319 602638232
20	440.2	32.1	456	13	BI966255 1E72904.y
21	436.6	31.4	440	13	BM128059 1E72904.y
22	431.6	31.4	538	12	BF821434 MRI-RF003
23	404	29.4	436	17	AO890280 HS_3188.B
24	399.4	29.1	413	9	BI762908 603047966
25	394	28.7	413	9	AI422796 tf22e06.x
26	389.2	28.3	698	13	BI906850 603064633
27	374.2	27.3	416	9	AI291866 qm6c02.x
28	372.6	27.1	584	10	AW917574 EST348878
29	365.4	26.6	422	13	BI677255 1E72904.y
30	365	26.6	399	9	AI913541 wa12c01.x
31	363.4	26.5	367	17	AO100365 HS_3054.A
32	359	26.1	407	9	BI965174 1E72904.y
33	358.4	26.1	407	9	AI221985 q955a10.x
34	357.6	26.0	894	13	BI908274 603068526
35	352.8	25.7	409	10	AW131279 xF61A04.x
36	347.2	25.3	397	9	AI669243 wc13901.x
37	342.6	25.0	710	12	BE858778 7E95005.x
38	341	24.8	377	13	BM509016 1H15005.y
39	339.8	24.7	372	13	BI677256 1H15005.y
40	330	24.0	374	13	BM505649 1H15005.x
41	328	23.9	367	13	BM662622 UT-E-CO1
42	327.8	23.9	650	12	BG404836 602420160
43	327	23.8	346	12	BE858822 7E95005.x
44	327	23.8	471	9	AA221610 my18d09.r
45	326	23.7	337	10	AW195034 x045912.x

## ALIGNMENTS

RESULT 1  
BI766766 834 bp mRNA linear EST 25-SEP-2001  
LOCUS 603056866F1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5406217 5',  
DEFINITION mRNA sequence.  
ACCESSION BI766766  
VERSION BI766766.1 GI:15758344  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 834)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: ggaaps-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L1AM1517 row: C column: 18  
High quality sequence stop: 772.  
Location/Qualifiers 1..834

FEATURES  
source



Db 237 ACCGCCAATCGGGAGATTATATGACACCCGGCTGGGCTACTACCTGATCTACTCAGG 296

Qy 605 TGCACCTTGTATGAGGGGAAAGCTGTCTACTCTAAGCTGGACTTGTGTGTGATGTGTGC 664

Db 297 TGCACCTTGTATGAGGGGAAAGCTGTCTACTCTAAGCTGGACTTGTGTGTGATGTGC 356

Qy 665 TGGCCCTCGCTGCTCGGAGAAATCTAGCACTGCGAGCCAGTTCCCTCGGGCCCCAGC 724

Db 357 TGGCCCTCGCTGCTCGGAGAAATCTAGCACTGCGAGCCAGTTCCCTCGGGCCCCAGC 416

Qy 725 TCCGCTCTGCCAGCTGTCTGGGGCTGTGGCCCTCGGGCCAGGATCTCCCTGCGGATCC 784

Db 417 TCCGCTCTGCCAGGCTGTCTGGGGCTGTGGCCCTCGGGCCAGGATCTCCCTGCGGATCC 476

Qy 785 GCACCTCCCGCTGGGCCCATCTCAGAGGCGCCCTCTCTCACTTCCAGCTTCC 844

Db 477 GCACCTCCCGCTGGGCCCATCTCAGAGGCGCCCTCTCTCACTTCCAGCTTCC 536

Qy 845 AGGTTCACTGAGGGGCGCCGTGTCTCCCAACAATCTCTCCAGGCTCCGAGTCCCTCGAC 904

Db 537 AGGTTCACTGAGGGGCGCCGTGTCTCCCAACAATCTCTCCAGGCTCCGAGTCCCTCGAC 596

Qy 905 AGCTCTGAGGAGACCGCGCCCTCTGGCCACCCCTCAGCGGCTTGTGTCTCAAA - CTT 963

Db 597 AGCTCTGAGGAGACCGCGCCCTCTGGCCACCCCTCAGCGGCTTGTGTCTCAAA CTT 656

Qy 964 GCCCCTCCCTAGAGGCTGCTCGGCGCTGTTCACGTGTTCATCCACATTAATACA 1023

Db 657 GCCCCTCCCTAGAGGCTGCTCGGCGCTGTTCACGTGTTCATCCACATTAATACA 716

Qy 1024 GATATCCACACTTATATTACAACTCCCCACCGCCCACTCTCCACCTTCACTAGCTCCG 1083

Db 717 GATATCCACACTTATATTACAAATCCCCACCGCAATCTCTCCACCTTCACTAGTTCGCG 776

Qy 1084 AATCCCTGACC - TTTAGAGGCCCCAGATGATCTGACATCCCCCTGGGCACACGCC 1141

Db 777 AATCCCTGACCCTTTAGAGGGCCCAAGGGAATCTCAACTCTCCCTCGGGCCAGAACCC 836

Qy 1142 AGGG 1145

Db 837 AAGG 840

RESULT 3	
BQ671259	
LOCUS	
DEFINITION	BQ671259                  963 bp        mRNA      linear   EST 15-JUL-2002 AGNCOURT_8303564 NIH_MGC_102 Homo sapiens CDNA IMAGE:62747716 5' , crna sequence.
ACCESSION	BQ671259
VERSION	BQ671259.1
KEYWORDS	GI:21782093
SOURCE	EST.
ORGANISM	human.  Homo sap lens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 963) NIH-MGC http://mgc.ncl.nih.gov/. National Institutes of Health. Mammalian Gene Collection (MGC) Unpublished (1999)
AUTHORS	Contact: Robert Strausberg, Ph.D.
JOURNAL	
COMMENT	

```

Plate: L1CM2456 row: 1 column: 13
High quality sequence stop: 565.
Location/Qualifiers
1..963
FEATURES
source
```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6274716"
/clone_id="NIH_MGC_102"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: salivary gland; Vector: pORF7; site_1: XhoI; site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT      194 a      326 c      260 g      183 t
ORIGIN
Query Match      54.4%; Score 747.2; DB 14; Length 963;
Best Local Similarity 91.9%; Pred. No. 9.5e-153;
Matches 868; Conservative 0; Mismatches 63; Indels 13; Gaps 7.

```

Query Match	54.4%	Score 747.2	DB 14	Length 963
Best Local Similarity	91.9%	Pred. No. 9.5e153		
Matches	868	Conservative	0	Mismatches 63
			Indels	13
			Gaps	7
QY 369	AGTTGGCGCTGCGAAGATGCACCTTAAAGGCGGAAAAACGSGGCTCGAAGACGATCGC	428		
DB 1	AGTTGGCGCTGCGAAGATGCACCTTAAAGGCGGAAAAACGSGGCTCGAAGACGATCGC	59		
QY 429	AGCCCATATATAGTTCATTCACGACCTGAGACGAGACGAGGCGACGACGATGTGACCG	488		
DB 60	AGCCCATATATAGTTCATTCACGACCTGAGACGAGACGAGGCGACGAGGATGTGACCG	119		
QY 489	GACAGTGAATGGCTGGGAGGAAGCAGATCAACAGCTCCACCCCTCTGCTACACCG	548		
DB 120	GACAGTGAATGGCTGGGAGGAAGCAGATCAACAGCTCCACCCCTCTGCTACACCG	179		
QY 549	CCAGATGGGAGATTTATATAGTACCCGGGCTGAGCTTACTACTGTACTGTACAGGTGA	608		
DB 180	CCAGATGGGAGATTTATATAGTACCCGGGCTGAGCTTACTACTGTACTGTACAGGTGA	239		
QY 609	CTTTGATGAGGGAAGGCTGTACTCTAAGGTGACTTGTGAGGGAATGAGTGTCTGGC	668		
DB 240	CTTTGATGAGGGAAGGCTGTACTCTAAGGTGACTTGTGAGGGAATGAGTGTCTGGC	299		
QY 669	CGTGGCTGCTGGAGGAATTTCTACGCCACTGCGGCGCCAGTTCCCTCGGCGCCACCTCGG	728		
DB 300	CGTGGCTGCTGGAGGAATTTCTACGCCACTGCGGCGCCAGTTCCCTCGGCGCCACCTCGG	359		
QY 729	CCTGTGCCAGGTGTCTGGGCTGTGTGGCCCTGCGGCGCAAGGATCTCTCCGATCCGAC	788		
DB 360	CCTGTGCCAGGTGTCTGGGCTGTGTGGCCCTGCGGCGCAAGGATCTCTCCGATCCGAC	419		
QY 789	CCCTCCCGTGGGCGCATTCACAGGCTGCGCCCTCTCCACCTACTCTGGAGCTCTTCCAGGT	848		
DB 420	CCCTCCCGTGGGCGCATTCACAGGCTGCGCCCTCTCCACCTACTCTGGAGCTCTTCCAGGT	479		
QY 849	TCACTGAGGGGCGCTGATCTCCCAACAGTGTGCCAGGCTGCGGCTCCCTCGACAGCT	908		
DB 480	TCACTGAGGGGCGCTGATCTCCCAACAGTGTGCCAGGCTGCGGCTCCCTCGACAGCT	539		
QY 909	CTCTGGGACACCCGGTCCCTCTGCGCCCAACCTCAGCCGCTCTTGTCTCAAGACGTGCCC	968		
DB 540	CTCTGGGACACCCGGTCCCTCTGCGCCCAACCTCAGCCGCTCTTGTCTCAAGACGTGCCC	599		
QY 969	TCCCTTAGAGGCTGCGGCGCTGTCACGATGTTTCCATCCCAATATAAGATATT	1028		
DB 600	TCCCTTAGAGGCTGCGGCGCTGTCACGATGTTTCCATCCCAATATAAGATATT	659		
QY 1029	CCCACTCTTATCTTACAACTCCCCC-ACCGCCACTCTCCACCTCAGTACTAGTCCCAATC	1087		
DB 660	CCCACTCTTATCTTACAACTCCCCC-ACCGCCCAACTCTCCACCTCAGTACTAGTCCCAAT	719		
QY 1088	CC-----TGACCTTTGAGAGGCCCCAGATATCTGCACTCCCCCTGCGCACAGACCCCGAG	1143		
DB 720	CCCGCTGACCCCTTTGAGAGGCCCCAGAGGATCTGCACTCCCCCTGCGCACAAACCCCGCA	779		

QY 1144 GCATT-CTGTCTACTCTGTGTGGCAAG-ATGGTCCAGAGACCCCACTTCAGG 1201  
 |||||  
 Db 780 GGCATTGGGTCTCACTGAATCTGGGGGAGAGAGGGGTCCAAAGACCCCTTTGGG 839  
 |||||  
 QY 1202 CACATAGAGGGGCT--GGACTGGCGGAGAGACCAAGAGACTGGG---CTTAGGCCA 1256  
 |||||  
 Db 840 ACTTAAAGGGGGCTTGACCTGGGGGAGAGAAACCCAAATAAATGGGGCCCAAGGCCA 899  
 |||||  
 QY 1257 GGAGTCCCAATGTGAGGGGCGAGAAACAAGACAGCTCTCC 1300  
 |||||  
 Db 900 GAAATTCACCAAGGAGGGGGGCGAATAAACAATAAGCTCC 943  
 |||||  
 RESULT 4 777 bp mRNA linear EST 04-OCT-2001  
 BI819200 603034614F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5175698 5',  
 LOCUS mRNA sequence.  
 DEFINITION  
 BI819200  
 VERSION BI819200.1 GI:15930750  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 777)  
 NIH-MGC http://mgc.nci.nih.gov/  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgabs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: L14M1437 row: 1 column: 03  
 High quality sequence stop: 759.  
 Location/Qualifiers  
 1..777  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_image="5175698"  
 /clone\_id="NIH\_MGC\_115"  
 /lab\_host="DH10B"  
 /note="Organ: pooled brain, lung, testis. Vector:  
 pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
 source anonymous pool of 6 male brains, age range 23-27; 1  
 male lung, age 27; and 1 male testis, age 69. Library is  
 oligo-dT primed and directionally cloned (EcoRV site is  
 destroyed upon cloning). Average insert size 1.8 kb,  
 insert size range 1-3 kb. Library is normalized and  
 enriched for full-length clones and was constructed by C.  
 Gruber (Invitrogen). Research Genetics tracking code  
 021. Note: this is a NIH\_MGC Library."  
 BASE COUNT 140 a 268 c 215 g 154 t  
 ORIGIN  
 Query Match 54.18; Score 743.4; DB 13; Length 777;  
 Best Local Similarity 99.18; Pred. No. 6.2e-132;  
 Matches 758; Conservative 0; Mismatches 6; Indels 1; Gaps 1;  
 QY 288 GGCAGAGAGAGACGACCGCTCGAATCCAGACAGAGAGAGAGAGATCC 347  
 |||||  
 Db 13 GGCAGAGAGAGACGACCGCTCGAATCCAGACAGAGAGAGAGAGATCC 72  
 |||||  
 QY 348 TGGCCCTTCTGACGACTAGTTCGGCTCGAGAGAGTGCACCTAAGGCGGAAAC 407  
 |||||  
 Db 73 TGGCCCTTCTGACGACTAGTTCGGCTCGAGAGAGTGCACCTAAGGCGGAAAC 132  
 |||||  
 QY 408 ACGGGCTCGAAGAGAGATCGACGCCATTATGAAGTTTCATCCAGACCTGACAGGCG 467  
 |||||

|||||  
 Db 133 ACGGGCTCGAAGAGAGATCGACGCCATTATGAAGTTTCATCCAGACCTGACAGGAGCG 192  
 |||||  
 QY 468 AGCCAGAGCAGGTGTGAGCGGAGAGATGATGCGTGGAGAGAACCAATCAACACTC 527  
 |||||  
 Db 193 AGCCAGAGCAGGTGTGAGCGGAGAGATGATGCGTGGAGAGAACCAATCAACACTC 252  
 |||||  
 QY 528 CAGCCCTCTGCGTACAAACCGCAGATCGGGAGTTTATGATCAACCGGCGCTGCTA 587  
 |||||  
 Db 253 CAGCCCTCTGCGTACAAACCGCAGATCGGGAGTTTATGATCAACCGGCGCTGCTA 312  
 |||||  
 QY 588 CTACCTGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 647  
 |||||  
 Db 313 CTACCTGACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 372  
 |||||  
 QY 648 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 707  
 |||||  
 Db 373 GCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 432  
 |||||  
 QY 708 TTCCCTGCGGCCAGCTCCGCTCTGCGAGGTGCTGGGCTGTTGGCCCTGCGGCGAG 767  
 |||||  
 Db 433 TTCCCTGCGGCCAGCTCCGCTCTGCGAGGTGCTGGGCTGTTGGCCCTGCGGCGAG 492  
 |||||  
 QY 768 GTCCCTGCGGCCAGCTCCGCTCTGCGAGGTGCTGGGCTGTTGGCCCTGCGGCGAG 827  
 |||||  
 Db 493 GTCCCTGCGGCCAGCTCCGCTCTGCGAGGTGCTGGGCTGTTGGCCCTGCGGCGAG 552  
 |||||  
 QY 828 CTACTGCGACTCTTCCAGGTTCACTGAGGGGCGGCTGCTCCACAGTGTCCGAGGC 887  
 |||||  
 Db 553 CTACTGCGACTCTTCCAGGTTCACTGAGGGGCGGCTGCTCCACAGTGTCCGAGGC 612  
 |||||  
 QY 888 TGGCGGCTCCCTCGACAGCTCTCTGCGACCGGCTCCCTCTGCTCCGACCGCTCAGCGC 947  
 |||||  
 Db 613 TGGCGGCTCCCTCGACAGCTCTCTGCGACCGGCTCCCTCTGCTCCGACCGCTCAGCGC 672  
 |||||  
 QY 948 TCTTTGCTCCAGACTGCGCTCCCTCTGAGAGGCTGCTGGGCTGTTACAGTGTTC 1007  
 |||||  
 Db 673 TCTTTGCTCCAGACTGCGCTCCCTCTGAGAGGCTGCTGGGCTGTTACAGTGTTC 732  
 |||||  
 QY 1008 ATCCCATATATACAGTATCCACATCTTCTTACAAATCC 1052  
 |||||  
 Db 733 ATCCCATATATACAGTATCCACATCTTCTTACAAATCC 776  
 |||||  
 RESULT 5 1071 bp mRNA linear EST 12-MAR-2002  
 BM921213 AGENCOURT-6633046 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5752561  
 LOCUS 5', mRNA sequence.  
 DEFINITION  
 BM921213  
 VERSION BM921213.1 GI:19371592  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 1071)  
 NIH-MGC http://mgc.nci.nih.gov/  
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
 TITLE Unpublished (1999)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgabs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: L14M12786 row: P column: 02  
 High quality sequence stop: 656.  
 Location/Qualifiers  
 1..1071

FEATURES  
 Source I..1071





OY	429	AGCCATTATGAAGTTCAATCCACGACCTGGACAGAGCGAGGCCAGCGAGGTGTGCACCG	488
Dd	241	AGCCCATTTATGAAGTTCATCACCAGACTTGAGAGAGAGCGAGCGACGAGTGTGCACCG	300
OY	489	GACAGTGAATGGCTGGAGGAGAACCCAGATCAACAGCTCCACCCCTCTGCGCTACAAACCG	548
Dd	301	GACAGTGAATGGCTGGAGGAGAACCCAGATCAACAGCTCCACCCCTCTGCGCTACAAACCG	360
OY	549	CCAGATCGGGGAGTTTATATAGTACCCCGGCTGGGCTCTACTACCTGTACTGTCAAGTGA	608
Dd	361	CCAGATCGGGGAGTTTATATAGTACCCCGGCTGGGCTCTACTACCTGTACTGTCAAGTGA	420
OY	609	CTTTATATGAGGGGAGGCGGTCTACTGTAAGCTGGACCTGACGTCGAGTAGAGTGTGCTGGC	668
Dd	421	CTTTATATGAGGGGAGGCGGTCTACTGTAAGCTGGACCTGACGTCGAGTAGAGTGTGCTGGC	480
OY	669	CCTGGCGTGCCTTGAGGAGATTTCTCACGCCACTGCGGCCAAGTTCCTTCGCGGCCACAGCTCG	728
Dd	481	CCTGGCGTGCCTTGAGGAGATTTCTCACGCCACTGCGGCCAAGTTCCTTCGCGGCCACAGCTCG	540
OY	729	CCTCTGCCACAGTGTCTGGGCTGTGGGCTGTGGCCCTGGGCGCCAGGCTCTCCCTGCGGATCCGCAC	788
Dd	541	CCTCTGCCACAGTGTCTGGGCTGTGGGCTGTGGCCCTGGGCGCCA-GGTCTCCCTGCGGATCCGCAC	599
OY	789	CTCTCCCTGGGCGCCATCTCAAGGCTGCCCCCTTCTCCACCTACTTCGAGCTCTTCCAGGT	848
Dd	600	CTCTCCCTGGGCGCCATCTCAAGGCTGCCCCCTTCTCCACCTACTTCGAGCTCTTCCAGGT	659
OY	849	TCACTGAGGGGGCCCTGATGCTCCCAACAGTGTCTCCACAGGCTGCGGCTCCCTCTGCACAGCT	908
Dd	660	TCACTGAGGGGGCCCTGATGCTCCCAACAGTGTCTCCCAACAGGCTGCGGCTCCCTCTGCACAGCT	719
OY	909	CTCTGGGCACC	919
Dd	720	CTCTGGGCACC	730

RESULT 7  
BI596681 LOCUS  
DEFINITION 603243254F1 NIH\_MGC\_96 Homo sapiens CDNA clone IMAGE:5285892 5', mRNA sequence.  
ACCESSION BI596681  
VERSION BI596681.1 GI:15489620  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 828)  
NIH-MGC http://mhc.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shitaki  
Toshiyuki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.lnl.gov  
Plate: LLM11722 row: k column: 13  
High quality sequence stop: 776.

```

/note="Organ: brain; Vector: pBluescriptplcr (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-xhoI (gtcagc
); Oligo-dn primed using primer 5'-tttttttttttttttvtvn-3'',
size-selected for average insert size 2.3 kb and
normalized to R0T 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NGRI, National
Institutes of Health). Note: this is a NIH_MGC Library."
155 a 278 c 223 g 172 t

```

Query Match	50.1%	Score 688	DB 13	Length 828
Best Local Similarity	95.3%	Pred. No. 7.8e-140		
Matches 732	Conservative	0	Mismatches 10	Indels 6
			Gaps	2
QY 312	GGAACTGAATCCCCAGACAGAAAGGACGAGATCTCGGCTTCTCTGAACCGACTAGT	371		
Db 43	GGAACTGAATCCCCAGACAGAAAGGACGAGATCTCGGCTTCTCTGAACCGACTAGT	102		
QY 372	TCGGCCTCGCAGAAAGTGCACCTAAAGGCCGGAAAAACAGGGGCTCGAAGAGGATCGACG	431		
Db 103	TCGGCCTCGCAGAAAGTGCACCTAAAGGCCGGAAAAACAGGGGCTCGAAGAGGATCGACG	162		
QY 432	CCATTATTAAGTCTATCCACGACCTCGGACAGGAGGAGGACGAGGAGGATGAGAGGAGAC	491		
Db 163	CCATTATTAAGTCTATCCACGACCTCGGACAGGAGGAGGAGGAGGAGGATGAGAGGAGAC	222		
QY 492	AGTAGTGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	551		
Db 223	ATTGAGTGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	282		
QY 552	GATCGGGAGGATTTATAGTACACCGGGGCTGGGCTCTACTACTGTACTGTACAGTGCACCTT	611		
Db 283	GATCGGGAGGATTTATAGTACACCGGGGCTGGGCTCTACTACTGTACTGTACAGTGCACCTT	342		
QY 612	TGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	671		
Db 343	TGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	402		
QY 672	GCGGCTGCTGGAGGAAATCTCAGCAGCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG	731		
Db 403	GCGGCTGCTGGAGGAAATCTCAGCAGCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG	462		
QY 732	CTGCCAGGTGCTGGGCTGTTGGCCCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	791		
Db 463	CTGCCAGGTGCTGGGCTGTTGGCCCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	522		
QY 792	CCCCGGGGCCCATCTCAAGAGGCTGCCCCCTTCTCTACACTTTCGAGACCTTCGAGGTTCA	851		
Db 523	CCCCGGGGCCCATCTCAAGAGGCTGCCCCCTTCTCTACACTTTCGAGACCTTCGAGGTTCA	582		
QY 852	CTGAGGGGCGCTGGGCTGCTCCCAAGTGTCTCCAGGCTGCGGAGGAGGAGGAGGAGGAGGAG	911		
Db 583	CTGAGGGGCGCTGGGCTGCTCCCAAGTGTCTCCAGGCTGCGGAGGAGGAGGAGGAGGAGGAG	642		
QY 912	TGGGACCCGGGTCCGCTTCGCCCCACCCCTCAGCGGCTTTTGTCTCCAGACCTGCCCCCTCC	971		
Db 643	TGGGACCCGGGTCCGCTTCGCCCCACCCCTCAGCGGCTTTTGTCTCCAGACCTGCCCCCTCC	702		
QY 972	CTTAGAGGCTGCGCTGGGCTGTTTCAGTGT---TTTCATCCCAATTAATACAGT-- 1025			
Db 703	CTTAGAGGCTGCGCTGGGCTGTTTCAGTGT---TTTCATCCCAATTAATACAGT-- 762			
QY 1026	ATTCCCACTCTTATCTTACACTCCCAACGCGCCACATCTCCACTCTCA 1073			
Db 763	TTTCCCACTCTTATCTTACACTCCCAACGCGCCACATCTCCACTCTCA 810			



fractionation: average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806. 0.5 microgram single stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an EcoT of 20. Single stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 142 a 128 c 209 g 129 t 1 others  
ORIGIN

Query Match 42.6%; Score 584.4; DB 13; Length 609;  
Best Local Similarity 99.7%; Pred. No. 3.1e-117;  
Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 787 ACCCTCCCTGGGCGCCATCTCAGAGGCTGCCCCCTTCTCAGCTACTTGGACTCTTCAG 846  
DB 609 ACCCTCCCTGGGCGCCATCTCAGAGGCTGCCCCCTTCTCAGCTACTTGGACTCTTCAG 550  
QY 847 GTTCACTAGAGGGGCGGCTGCTCCCAAGTCTCCAGAGGCTGCGGGCTCCCTCGACAG 906  
DB 549 GTTCACTAGAGGGGCGGCTGCTCCCAAGTCTCCAGAGGCTGCGGGCTCCCTCGACAG 490  
QY 907 CTCTCTGGGACACCGGCTCCCTCTGCCCCACCTCAGCCGCTCTTCTCTCAGACCTGCG 966  
DB 489 CTCTCTGGGACACCGGCTCCCTCTGCCCCACCTCAGCCGCTCTTCTCTCAGACCTGCG 430  
QY 967 CTTCCCTCTAGAGGCTGCTGCGGGCTGTTACAGTCTTTCATCCACATTAATACAGTA 1026  
DB 429 CTTCCCTCTAGAGGCTGCTGCGGGCTGTTACAGTCTTTCATCCACATTAATACAGTA 370  
QY 1027 TTCCCACTCTTATCTTAACTCCCAACGCGCCACTCTCAGCTACTAGCTCCCAAT 1086  
DB 369 TTCCCACTCTTATCTTAACTCCCAACGCGCCACTCTCAGCTACTAGCTCCCAAT 310  
QY 1087 CCTCTGACCTTTTGAAGGCGCCAGTGAATCTGACTCCCTGCGCACAGACCCCGAGGC 1146  
DB 309 CCTCTGACCTTTTGAAGGCGCCAGTGAATCTGACTCCCTGCGCACAGACCCCGAGGC 250  
QY 1147 ATTTGTCTACTGACTCTGTTGGGCAAGAGTGGTCCAGAAAGACCCCACTTCAGGCACTA 1206  
DB 249 ATTTGTCTACTGACTCTGTTGGGCAAGAGTGGTCCAGAAAGACCCCACTTCAGGCACTA 190  
QY 1207 AGAGGGGCTGAGCTGCGGCGAGGAAAGCAAGAGACTGCGGCTTAGGCAAGATTCCTCA 1286  
DB 189 AGAGGGGCTGAGCTGCGGCGAGGAAAGCAAGAGACTGCGGCTTAGGCAAGATTCCTCA 130  
QY 1267 AATGTAGGGGCGGAAACAAGACAGCTCTCCCTTGAGAAATCCCTGTGGATTTTAA 1326  
DB 129 AATGTAGGGGCGGAAACAAGACAGCTCTCCCTTGAGAAATCCCTGTGGATTTTAA 70  
QY 1327 AACGATATATTTTATTTATTTATTTGACAAATGTTGATTAATGG 1373  
DB 69 AACGATATATTTTATTTATTTATTTGACAAATGTTGATTAATGG 23

RESULT 10  
AF163779 1027 bp DNA linear GSS 29-AUG-2000  
LOCUS AF163779 Human Homo sapiens genomic clone BAC750E14, DNA sequence.  
DEFINITION AF163779  
ACCESSION AF163779  
VERSION AF163779.1 GI:5726439  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
JOURNAL Cousin, P., Billotte, J., Chabert, P., and Shaw, P. H.  
Physical map of 17p13 and the genes adjacent to p53  
Genomics 63 (1), 60-66 (2000)

MEDLINE 20130114  
COMMENT

Contact: Shaw PH  
Experimental Oncology  
Institute of Pathology  
Rue du Bugnon 25, Lausanne, VD 1011, Switzerland  
Subclone-A2R Asc-BamHI PSL180  
Class: BAC subclone.  
location/Qualifiers  
1. 1027  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="17p"  
/clone="BAC750E14"  
/clone\_lib="Human"  
/note="Vector: PMOS Blue"

FEATURES  
source

BASE COUNT 207 a 317 c 282 g 208 t 13 others  
ORIGIN

Query Match 42.3%; Score 581.4; DB 17; Length 1027;  
Best Local Similarity 98.0%; Pred. No. 1.5e-116;  
Matches 577; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

QY 785 GCACCTTCCCTGGGCGCCATCTCAGAGGCTGCCCCCTTCTCAGCTACTTGGACTCTTC 844  
DB 1 GCWCCCTCCCTGGGCGCCATCTCAGAGGCTGCCCCCTTCTCAGCTACTTGGACTCTTC 60  
QY 845 AGTTCACTAGAGGGGCGGCTGCTCCCAAGTCTCCAGAGGCTGCGGGCTCCCTCGAC 904  
DB 61 AGTTCACTAGAGGGGCGGCTGCTCCCAAGTCTCCAGAGGCTGCGGGCTCCCTCGAC 120  
QY 905 AGCTCTGAGGACACCGGCTCCCTCTGCCCCACCTCAGCCGCTCTTGTCTCAGAGCTG 964  
DB 121 AGCTCTGAGGACACCGGCTCCCTCTGCCCCACCTCAGCCGCTCTTGTCTCAGAGCTG 180  
QY 965 CCCCTCCCTAGAGGCTGCTGCGGGCTGTTACAGTCTTTCATCCACATTAATACAG 1024  
DB 181 CCCCTCCCTAGAGGCTGCTGCGGGCTGTTACAGTCTTTCATCCACATTAATACAG 240  
QY 1025 TATTCACACTTATCTTACAACTCCCAACGCGCCACTCTCAGCTACTAGCTCCCA 1084  
DB 241 TATTCACACTTATCTTACAACTCCCAACGCGCCACTCTCAGCTACTAGCTCCCA 300  
QY 1085 ATCCCTACCTTTTGAAGGCGCCAGTGAATCTGACTCCCTGCGCACAGACCCCGAGG 1144  
DB 301 ATCCCTACCTTTTGAAGGCGCCAGTGAATCTGACTCCCTGCGCACAGACCCCGAGG 360  
QY 1145 GCATTGTCTACTGACTCTGTTGGGCAAGAGTGGTCCAGAAAGACCCCACTTCAGGCACT 1204  
DB 361 GCATTGTCTACTGACTCTGTTGGGCAAGAGTGGTCCAGAAAGACCCCACTTCAGGCACT 420  
QY 1205 TAAGAGGGGCTGAGCTGCGGCGAGGAAAGCAAGAGACTGCGGCTTAGGCAAGATTCCTCA 1264  
DB 421 TAAGAGGGGCTGAGCTGCGGCGAGGAAAGCAAGAGACTGCGGCTTAGGCAAGATTCCTCA 480  
QY 1265 CAAATGTAGAGGGGCGGAAACAAGACAGCTCTCCCTTGAGAAATCCCTGTGGATTTT 1324  
DB 481 CAAATGTAGAGGGGCGGAAACAAGACAGCTCTCCCTTGAGAAATCCCTGTGGATTTT 540  
QY 1325 AAAACGATATTTTATTTATTTATTTATTTGACAAATGTTGATTAATGG 1373  
DB 541 AAAACGATATTTTATTTATTTATTTATTTGACAAATGTTGATTAATGG 589

RESULT 11  
BO884231 940 bp mRNA linear EST 16-AUG-2002  
LOCUS BO884231  
DEFINITION AGENCOURT\_8682031 lupsk1\_sclatlc\_nerve Homo sapiens cdna clone  
IMAGE:6197488 5', mRNA sequence.  
ACCESSION BO884231  
VERSION BO884231.1 GI:22276239  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 940)  
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
 NIH-MGC http://mgc.ncl.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1998)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: Dr. James R. Lupski  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: L1AM13607 row: 3 column: 17  
 High quality sequence stop: 453.

## FEATURES

Location/Qualifiers  
 1..940  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6197488"  
 /clone\_lib="Lupski\_sciatic\_nerve"  
 /sex="male"  
 /tissue\_type="sciatic nerve"  
 /dev\_stage="adult, 70 yr"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site\_1:  
 NotI; Site\_2: SalI; CDNA made by oligo-dT priming.  
 Directionally cloned using the following adaptors:  
 5'-TCGACCCACGCGCCG-3' and  
 5'-GACTAGTCTGATGATGAGCGCGCCCTT(15'-3', Size selected >  
 1 kb for average insert length 1.87 kb. This is a primary  
 library, non-amplified. Library constructed by Life  
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
 College of Medicine) and is available through Life  
 Technologies."

BASE COUNT 170 a 325 c 255 g 190 t  
 ORIGIN

Query Match 40.9%; Score 562; DB 14; Length 940;  
 Best Local Similarity 96.5%; Pred. No. 2.6e-112;  
 Matches 660; Conservative 0; Mismatches 15; Indels 9; Gaps 8;

301 CAGGACCCGTCGGAAGTATCCAGAGAGAAAGCAGGATCTGGCCCTTCTCG 360  
 1 CAGGACCCGTCGGAAGTATCCAGAGAGAAAGCAGGATCTGGCCCTTCTCG 60  
 361 AACGCACTAGTTCGGCTCGCAGAGAGTCCACCTAAAGCCGGAACACAGGCTCGAAGA 420  
 61 AACGCACTAGTTCGGCTCGCAGAGAGTCCACCTAAAGCCGGAACACAGGCTCGAAGA 120  
 421 GCGATCGAGCCCATTAAGAGTTTATCAGACCTGAGACAGGAGCGAGCAGCAGT 480  
 121 GCGATCGAGCCCATTAAGAGTTTATCAGACCTGAGACAGGAGCGAGCAGCAGT 180  
 481 GTGACGGGACAGTGTGCTGGAGGAGAGCAGATCAACAGTCCAGGCCCTTGCGC 540  
 181 GTGACGGGACAGTGTGCTGGAGGAGAGCAGATCAACAGTCCAGGCCCTTGCGC 240  
 541 TACAACCCGACAGTGGGAGATTATAGTCAACCCGGGCTGAGCTTACTACTACTGT 600  
 241 TACAACCCGACAGTGGGAGATTATAGTCAACCCGGGCTGAGCTTACTACTACTGT 300  
 601 CAGTGCACCTTTGATGAGAGGAGGCTCTTACCTGAAGCTTGAGCTTGTGATGATGT 660  
 301 CAGTGCACCTTTGATGAGAGGAGGCTCTTACCTGAAGCTTGAGCTTGTGATGATGT 360  
 661 GTGCTGGGCTGCGCTGCTGAGAGATTCACAGCACTGGGCGCAGTTCCCTGGGGCC 720  
 361 GTGCTGGGCTGCGCTGCTGAGAGATTCACAGCACTGGGCGCAGTTCCCTGGGGCC 420  
 721 CAGTCCGCGCTGTG-CGAGTGTGTGGCTGTGGCCCTGGGGCA-GGTCCTTCCCTGC 778

421 CAGTCCGCGCTGTGCGCAGAGTGTGCGGCTGTGGCCCTGGGGCCAGGGCCCTTCCCTGC 480  
 779 GGATCCGACACCTCCCTCGGGCCCATCTCAGAGCTG-CGCCCTTCTCAGCTAC-TTCGG 836  
 481 GGATCCGACACCTCCCTCGGGCCCATCTCAGAGCTGCGCCCTTCTCAGCTACTTTCGG 540  
 837 ACTCTCCAGGTTCACTAGAGGGCCCTGCTGCCACAGTCGT-CCAGGCTGCGGGC- 894  
 541 ACTCTCCAGGTTCACTAGAGGGCCCTGCTGCCACAGTCGTCCACAGGCTGCGGGC 600  
 895 TCCCTCCGACAGCTCTCT-GGGCACCAGGCTCCCTCTGCCACACC--TCAGCGCTCTT 951  
 601 TCCTCCGACAGCTCTCTGGGGAACCGGTCCCTCTGCCACACCCTCAGCGGCTTCTT 660  
 952 TGCTCCAGAGCTGCGCCCTCTCT 975  
 661 TGCTCCAGAGCTGCGCCCTCTCT 684

RESULT 12  
 BM971606/c 568 bp mRNA linear EST 21-MAR-2002  
 LOCUS  
 DEFINITION UI-CF-EC1-ab1-p-06-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone  
 UI-CF-EC1-ab1-p-06-0-UI 3', mRNA sequence.  
 ACCESSION BM971606  
 VERSION  
 KEYWORDS BM971606.1 GI:19589193  
 EST.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 568)  
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 COMMENT Contact: McCray, PB  
 McCray Lab  
 University of Iowa  
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
 Tel: 319 356 4866  
 Fax: 319 356 7171

Email: paul-mccray@uiowa.edu  
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com)  
 The following repetitive elements were found in this CDNA  
 sequence: 1-82, >AT-rich/low-complexity (matched complement)  
 Seq primer: M13 FORWARD  
 POLYA=yes.

## FEATURES

Location/Qualifiers  
 1..568  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="UI-CF-EC1-ab1-p-06-0-UI"  
 /clone\_lib="UI-CF-EC1"  
 /tissue\_type="Lung"  
 /dev\_stage="Adult and Fetal"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /note="Organ: Lung; Vector: pVT3-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: BcoR I; Site\_2: Not I;  
 UI-CF-EC1 is a normalized CDNA library containing the  
 following tissue(s): Normal lung from adult and from fetal  
 day 64, day 87, week 19 and week 42. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand CDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded CDNA was ligated to an BcoR I

adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is AAGTCTTAC.

TAG: LIB-UI-CF-ECL  
TAG: Tissue=Normal Lung Epithelial Cells Tissue nos 369-371 and 380-383  
TAG\_SEQ=AACTGCTTAC"

BASE COUNT 130 a 119 c 184 g 135 t  
ORIGIN

Query Match 38.9%; Score 534.6; DB 14; Length 568;  
Best Local Similarity 99.3%; Pred. No. 2.3e-106;  
Matches 537; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

833 TCGGACTCTTCAGGTTACTAGAGGGCCCTGCTCCCAACAGTGTCCAGGCTGGCG 892  
DB TCGGACTCTTCAGGTTACTAGAGGGCCCTGCTCCCAACAGTGTCCAGGCTGGCG 509  
893 GCTCCCTTCAGAGCTCTCTGGGACCCGGTCCCTTGCACCCTCAGCCGCTCTTT 952  
DB GCTCCCTTCAGAGCTCTCTGGGACCCGGTCCCTTGCACCCTCAGCCGCTCTTT 449  
953 GCTCCAGAGCTGCCCTCCCTCTAGAGGCTGCTGGGCTGTACAGTGTTCATCCC 1012  
DB GCTCCAGAGCTGCCCTCCCTCTAGAGGCTGCTGGGCTGTACAGTGTTCATCCC 389  
1013 ACATTAATACAGTATTTCCACTCTTATCTTACACTCCCAACGCGCCACTCTCCACCTC 1072  
DB ACATTAATACAGTATTTCCACTCTTATCTTACACTCCCAACGCGCCACTCTCCACCTC 329  
1073 ACTAGCTCCCAATCCCTGACACCTTTGAGGGCCCAAGTGTGATCTGCATCCCTGCGCA 1132  
DB ACTAGCTCCCAATCCCTGACACCTTTGAGGGCCCAAGTGTGATCTGCATCCCTGCGCA 269  
1133 CAGACCCCAAGGAGTGTCTTCACTCTGTGAGGCAAGATGGTCCAGAAAGACC 1192  
DB CAGACCCCAAGGAGTGTCTTCACTCTGTGAGGCAAGATGGTCCAGAAAGACC 209  
1193 CACTTCAGGACTTAAGAGGGGCTGACCTGGCGGAGAAACCAAGAGTGGGCTTAG 1252  
DB CACTTCAGGACTTAAGAGGGGCTGACCTGGCGGAGAAACCAAGAGTGGGCTTAG 149  
1253 GCCAGAGTTCGCAAAATGTGAGGGGCGAGAAACAAGACACTCTCTCTGAGAATTC 1312  
DB GCCAGAGTTCGCAAAATGTGAGGGGCGAGAAACAAGACACTCTCTCTGAGAATTC 89  
1313 CTGTGATTTTAAACAGATATTTATTTATTTATTTATTTGACAAATGTTGATTAATG 1372  
DB CTGTGATTTTAAACAGATATTTATTTATTTATTTATTTATTTGACAAATGTTGATTAATG 29  
1373 G 1373  
DB 28 g 28

RESULT 13  
LOCUS Bg110063 910 bp mRNA linear EST 30-JAN-2001  
DEFINITION 602279667F1 NIH\_MGC\_86 Homo sapiens cDNA clone IMAGE:4367225 5',  
mRNA sequence.  
ACCESSION Bg110063  
VERSION Bg110063.1 GI:12603569  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 910)  
AUTHORS NIH-MGC <http://mgc.ncl.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@b-remail.nih.gov  
Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.lnl.gov>  
Plate: L1AM10019 row: e column: 18  
High quality sequence stop: 493.

## FEATURES

source

Location/Qualifiers  
1..910  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4367225"  
/clone\_1id="NIH-MGC\_86"  
/tissue\_type="Osteosarcoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: bone; Vector: pCMV-SPORT6; Site:1; Noti:  
Site:2; SalI: Cloned unidirectionally; oligo-dr primed.  
Average insert size 1.53 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH-MGC library."

BASE COUNT 262 a 279 c 225 g 144 t  
ORIGIN

Query Match 38.7%; Score 531.6; DB 12; Length 910;  
Best Local Similarity 88.6%; Pred. No. 1.1e-105;  
Matches 622; Conservative 0; Mismatches 74; Indels 6; Gaps 4;

627 TGTCTACCTGAAGCTGAGCTGTGTGATGATGATGCTGCGCTGCGCTGAGAGA 686  
DB 1 TGTCTACCTGAAGCTGAGCTGTGTGATGATGATGCTGCGCTGCGCTGAGAGA 60  
687 ATTCTCAGCACTGCGGCACTTCCCTGCGGCCCAAGCTCCGCTTCCAGGTGTGG 746  
DB ATTCTCAGCACTGCGGCACTTCCCTGCGGCCCAAGCTCCGCTTCCAGGTGTGG 120  
61 ATTCTCAGCACTGCGGCACTTCCCTGCGGCCCAAGCTCCGCTTCCAGGTGTGG 120  
747 GCTGTGGCCCTGCGGCAAGGCTCCTCCCTCGGATCGGACCCCTCCGCGGCAATCT 806  
DB GCTGTGGCCCTGCGGCAAGGCTCCTCCCTCGGATCGGACCCCTCCGCGGCAATCT 179  
121 GCTGTGGCCCTGCGGCAAGGCTCCTCCCTCGGATCGGACCCCTCCGCGGCAATCT 179  
807 CAAGCTGCCCCCTTCTCTACACTTCTGCACTCTTCCAGGTTCACTGAGGGGCTGTGT 866  
DB CAAGCTGCCCCCTTCTCTACACTTCTGCACTCTTCCAGGTTCACTGAGGGGCTGTGT 239  
180 CAAGCTGCCCCCTTCTCTACACTTCTGCACTCTTCCAGGTTCACTGAGGGGCTGTGT 239  
867 CTCCCAACAGTGTGCTCCCAAGGCTGCGGCTCCCTCGACAGCTCTGTGGCAACCGGTGCC 926  
DB CTCCCAACAGTGTGCTCCCAAGGCTGCGGCTCCCTCGACAGCTCTGTGGCAACCGGTGCC 299  
240 CTCCCAACAGTGTGCTCCCAAGGCTGCGGCTCCCTCGACAGCTCTGTGGCAACCGGTGCC 299  
927 CTGTGCCCAACCTTCAGCGGCTTGTGCTCCAGACGCGCCCTCCCTGAGAGGCTGCT 986  
DB CTGTGCCCAACCTTCAGCGGCTTGTGCTCCAGACGCGCCCTCCCTGAGAGGCTGCT 359  
300 CTGTGCCCAACCTTCAGCGGCTTGTGCTCCAGACGCGCCCTCCCTGAGAGGCTGCT 359  
987 GGGCTGTTCAGCTGTTTTCATCCCAATTAATACAGTATTTCCACTTATCTTACAA 1046  
DB GGGCTGTTCAGCTGTTTTCATCCCAATTAATACAGTATTTCCACTTATCTTACAA 419  
360 GGGCTGTTCAGCTGTTTTCATCCCAATTAATACAGTATTTCCACTTATCTTACAA 419  
1047 CTCCCCACAGGCGCACTCTCAGCTACCTAGCTCCCAATTCCTGACCTTTGAGGCCCC 1106  
DB CTCCCCACAGGCGCACTCTCAGCTACCTAGCTCCCAATTCCTGACCTTTGAGGCCCC 478  
420 CTCCCCACAGGCGCACTCTCAGCTACCTAGCTCCCAATTCCTGACCTTTGAGGCCCC 478  
1107 CAGTATCTCGACTCCCGCTTGGGCAAGAGCCCGGAGGATGTGTCTACTCTG 1166  
DB CAGTATCTCGACTCCCGCTTGGGCAAGAGCCCGGAGGATGTGTCTACTCTG 538  
479 AGTGATCTCGATCCCGCACTGACACAGCCCGGAGGATGTGTCTACTCTG 538  
1167 TGGGCAAGATGGGTCCAGAGAGCCCACTTCAAGCACTTAAGAGGCTGAGCTGGCG 1226  
DB TGGGCAAGATGGGTCCAGAGAGCCCACTTCAAGCACTTAAGAGGCTGAGCTGGCG 595  
539 GAGGCAAGATGGGTCCAGAGAGCCCACTTCAAGCACTTAAGAGGCTGAGCTGGCG 595  
1227 CAGGAAGCCAAAGAGACTG-GGCTTAGGCCAGAGGTTCCCAATGTGAGGGGCGAGAAAC 1285

QY	1286	AAGACAAGCTCTCTCCCTTGAGAAATTCCTGTGGATTTTAA	1327
Db	596	AGGGAAGCAAGAGACTGAGGCTGAGGCAAGACAGATTTCCAAAGTGAAGGCGCAAGAA	655
QY	1286	AAGACAAGCTCTCTCCCTTGAGAAATTCCTGTGGATTTTAA	1327
Db	656	CAAGGAAGACCTCCATGAGGATGAGTGGCCTGAGAAATATGACAA	697
RESULT 14			
LOCUS	BI824443	531 bp	mRNA
DEFINITION	603038693p1 NIH_MGC_115	Homo sapiens	cdna clone IMAGE:5179510 5',
ACCESSION	BI824443		
VERSION	BI824443.1	GI:15935993	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.ncl.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strusberg, Ph.D.		
	Email: cga@phs-remail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LMNL at:		
	http://image.lnl.gov		
	Plate: LMNL1447 row: j column: 23		
FEATURES	High quality sequence stop: 529.		
SOURCE	Location/Qualifiers		
	1..531		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone_image="5179510"		
	/clone_11b="NIH_MGC_115"		
	/lab_host="DH10B"		
	/note="Organ: pooled brain, lung, testis; Vector: pcwv-sport6; site_1: NotI; site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."		
BASE COUNT	72 a	223 c	119 g
ORIGIN			117 t
Query Match	37.6%	Score 516.8;	DB 13; Length 531;
Best Local Similarity	99.4%;	Prod. No. 1.7e-102;	
Matches 529; Conservative	0;	Mismatches 2;	Indels 1; Gaps 1;
QY	661	GTGCTGAGCCCGGCTGCTGAGGAATTCACGACATCGGCGCAGTTCCCTCGGGCC	720
Db	1	GTGCTGAGCCCGGCTGCTGAGGAATTCACGACATCGGCGCAGTTCCCTCGGGCC	60
QY	721	CAGTCCCGCTCTGCTGAGGATGCTGGGCTGTGGCCCTGCGGCGAGGCTCCCTCGGG	780
Db	61	CAGTCCCGCTCTGCTGAGGATGCTGGGCTGTGGCCCTGCGGCGAGGCTCCCTCGGG	119
QY	781	ATTCGACACCTTCCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT	840
Db	120	ATTCGACACCTTCCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT	179
QY	841	TTTCAGGTTACTAGAGGCTTGGTCTGCTCCCAAGTGTCCAGGCTGCGGCTCCCT	900

[illegible]







(UYGE-) UNIV GENEVA FACULTY MEDICINE.

XX Browning JL, Chicopeportliche Y;

XX MPI: 1998-145619/13.

DR P-PSDB; AAM47525.

PT Tumour necrosis factor related ligand - useful for, e.g. treating  
PT cancer, auto-immune disease and immune responses to tissue grafts  
XX  
XX Claim 2; Pages 48-50; 69pp; English.

The sequence is that encoding human tumour necrosis factor related  
ligand (TRELL). TRELL or active fragments can be included with a  
carrier in pharmaceutical compositions to treat cancer, autoimmune  
diseases or immune responses to tissue grafts, or to stimulate or  
suppress the immune system. It is useful to screen for TRELL  
receptors, by labelling with a detectable label and screening  
compositions for binding. Agents interfering with TRELL-receptor  
binding can also be screened for, can then be administered,  
optionally with interferon- gamma, to induce cell death or  
treat, suppress or alter immune responses (especially involving human  
adenocarcinoma cells) involving a signal pathway between TRELL and its  
receptor. The DNA sequence can be used in gene therapy for  
CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
CC autoimmune and inflammatory diseases or inherited genetic disorders,  
CC by introducing into cells, and expressing, therapeutically effective  
amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
CC It may also be of use in the preparation of probe probes for  
CC screening natural/synthetic DNAs for TRELL-encoding sequences  
CC and for antisense therapy.

XX Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other;

Query Match 100.0%; Score 1373; DB 19; Length 1373;  
Best Local Similarity 100.0%; Pred. No. 7e-283;  
Matches 1373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCATTTGTAAGCTTTGAATTTCCGCGCGCTCCCGCTCCCGCATCTCTGGG 60  
DB 1 ATGTCATTTGTAAGCTTTGAATTTCCGCGCGCTCCCGCTCCCGCATCTCTGGG 60  
QY 61 TCCCGGATGGGGGGGGGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
DB 61 TCCCGGATGGGGGGGGGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
QY 121 AGCCAGAGGGGAGGGGGGGGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
DB 121 AGCCAGAGGGGAGGGGGGGGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180  
QY 181 CTGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG 240  
DB 181 CTGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG 240  
QY 241 AGCCGGGATGCTGTCGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
DB 241 AGCCGGGATGCTGTCGGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300  
QY 301 CAGGACCGCTGCAATGTAATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
DB 301 CAGGACCGCTGCAATGTAATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360  
QY 361 AACCGACTAGTGGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
DB 361 AACCGACTAGTGGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420  
QY 421 GCGATGCGACGCAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 480  
DB 421 GCGATGCGACGCAATTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAAT 480  
QY 481 GTGGAGCGGAG 540  
DB 481 GTGGAGCGGAG 540

QY 541 TACACCGCCAGATGGGGAGATTATATGACACCGGGCTGGGCTACTACTGTAAGT 600  
DB 541 TACACCGCCAGATGGGGAGATTATATGACACCGGGCTGGGCTACTACTGTAAGT 600  
QY 601 CAGTGCATCTTGTATGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
DB 601 CAGTGCATCTTGTATGAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660  
QY 661 GTGCTGGGCGGCTGGGCTGGGAGAAATTCACGACAGTGGGAGAGTCCCTGGGCGC 720  
DB 661 GTGCTGGGCGGCTGGGCTGGGAGAAATTCACGACAGTGGGAGAGTCCCTGGGCGC 720  
QY 721 CAGCTCGGCTCTGCAAGAGTGTGGCTGTGGCCCTGGGCGAGAGAGTCCCTGGC 780  
DB 721 CAGCTCGGCTCTGCAAGAGTGTGGCTGTGGCCCTGGGCGAGAGAGTCCCTGGC 780  
QY 781 ATCCGACACCTCCCTGGGCGCAATCTAAGGCTCCCTCTCTACTACTTGGAGTC 840  
DB 781 ATCCGACACCTCCCTGGGCGCAATCTAAGGCTCCCTCTCTACTACTTGGAGTC 840  
QY 841 TTCAGGTTCACTAGAGGGGCTGTGCTCCCAACAGTCTGCCAGGCTGGCGGCTCCCT 900  
DB 841 TTCAGGTTCACTAGAGGGGCTGTGCTCCCAACAGTCTGCCAGGCTGGCGGCTCCCT 900  
QY 901 CGACAGCTCTGTGGGAGACCGGCTCCCTGTGCGCCACCTCAGGCGCTTTGGTCCAGA 960  
DB 901 CGACAGCTCTGTGGGAGACCGGCTCCCTGTGCGCCACCTCAGGCGCTTTGGTCCAGA 960  
QY 961 CCGTCCGCTCCCTTAAGAGTGTGGGCTGTGCTCCCAAGTCTTCCATCCCATTAAT 1020  
DB 961 CCGTCCGCTCCCTTAAGAGTGTGGGCTGTGCTCCCAAGTCTTCCATCCCATTAAT 1020  
QY 1021 ACAGATTTCCCACTTATATCTTACAACTCCCGCCAGCCAGCTCCCACTACTAGCTC 1080  
DB 1021 ACAGATTTCCCACTTATATCTTACAACTCCCGCCAGCCAGCTCCCACTACTAGCTC 1080  
QY 1081 CCCATCTCTGACCTTTGAGGCGCCAGTATCTGACATCCCTCTGGCCACAGACCC 1140  
DB 1081 CCCATCTCTGACCTTTGAGGCGCCAGTATCTGACATCCCTCTGGCCACAGACCC 1140  
QY 1141 CAGGCAATGTTGTCATGACTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
DB 1141 CAGGCAATGTTGTCATGACTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200  
QY 1201 GCACATAAGAGGGGCTGAGACCTGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
DB 1201 GCACATAAGAGGGGCTGAGACCTGGGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 1260  
QY 1261 TTCCCAATGTGAGGGGCGAGAAACAAGACAGCTCCCTTGAAGATTCCTGTGGAT 1320  
DB 1261 TTCCCAATGTGAGGGGCGAGAAACAAGACAGCTCCCTTGAAGATTCCTGTGGAT 1320  
QY 1321 TTTTAAACAGATTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1373  
DB 1321 TTTTAAACAGATTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1373

RESULT 2  
ABK34881  
ID ABK34881 standard; cDNA; 1364 BP.  
XX ABK34881;  
AC  
XX  
XX 08-MAY-2002 (first entry)  
XX  
XX Human cDNA encoding secreted protein #19.  
XX  
XX Human; secreted protein; gene; ss; nutritional supplement; haemophilia;  
KW viral infection; bacterial infection; fungal infection; diabetes; asthma;  
KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;  
KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;  
KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;

KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;  
KW tissue regeneration; wound healing; burn; hematopoiesis;  
KW myeloid cell deficiency; lymphoid cell deficiency.

XX Homo sapiens.

XX WO200177288-A2.

XX 18-OCT-2001.

XX 29-MAR-2001; 2001WO-US10224.

XX 06-APR-2000; 2000US-195582P.

XX (GEMV ) GENETICS INST INC.

XX Wong GG, Clark HF, Fiechtel K, Agostino MJ, Howes SH, Resnick RJ,  
PI Gulikota K, Graham JR;

XX WPI: 2002-179321/23.

XX Five hundred and ninety two polynucleotides derived from a variety of  
PT human tissue sources which encode secreted proteins, useful for  
PT treating immune deficiencies and disorders such as autoimmune disorders

PS Claim 1; Page 82; 372pp; English.

CC The invention relates to 592 polynucleotides which have been derived from  
CC a variety of human tissue sources and which encode novel secreted  
CC proteins. The polynucleotides can be used as probes for the  
CC identification and isolation of full length cDNA and genomic DNA. The  
CC polynucleotides and proteins can also be used as nutritional supplements.  
CC The proteins are useful in the treatment of various immune deficiencies  
CC and disorders such as viral infections, bacterial infections, fungal  
CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple  
CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions  
CC and conditions (e.g. asthma). They are also useful for treating  
CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's  
CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),  
CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also  
CC useful for tissue regeneration, for wound healing and in the treatment of  
CC burns, incisions and ulcers. The proteins are also useful for regulating  
CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.  
CC Sequences ABK34863-ABK3454 represent polynucleotides of the invention.  
XX

XX Sequence 1364 BP; 246 A; 461 C; 394 G; 263 T; 0 other;

Query Match 98.0%; Score 1345.8; DB 24; Length 1364;

Best Local Similarity 99.9%; Pred. No. 4.3e-277;  
Matches 1347; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 25 TCCGCCCGCGGCTCCCTCCCGATCCCTCGGCTCCCGGATGGGGGCGGTGAGG 84  
DB 5 TCCGCCCGCGGCTCCCTCCCGATCCCTCGGCTCCCGGATGGGGGCGGTGAGG 64  
QY 85 CAGGCACAGCCCGCGCGCCCATGGGCGCCGTCGAGCAGAGGGGGGCGGCCGG 144  
DB 65 CAGGCACAGCCCGCGCGCCCATGGGCGCCGTCGAGCAGAGGGGGGCGGCCGG 124  
QY 145 GGGGAGCCCGGACCGCTCTGCTGTCCTCGCTCGCGCTGGGCTGGGCTGGGCTGGCC 204  
DB 125 GGGGAGCCCGGACCGCTCTGCTGTCCTCGCTCGCGCTGGGCTGGGCTGGGCTGGCC 184  
QY 205 TGCCTTGCGCTCTGCTGTCCTGCTGTCATGTTGGGGAGCGCGCATGCTGTCGCCAG 264  
DB 185 TGCCTTGCGCTCTGCTGTCCTGCTGTCATGTTGGGGAGCGCGCATGCTGTCGCCAG 244  
QY 265 GAGCCGCGCCAGGAGAGTGTGGCAGAGGAGGAGCAGAGCCCGCGCATGATGATGCC 324  
DB 245 GAGCCGCGCCAGGAGAGTGTGGCAGAGGAGGAGCAGAGCCCGCGCATGATGATGCC 304  
QY 325 CAGACAGAAAGAACCGAGATCTGCGCTTCTGTAACGACTAGTGGCGCTCGAGA 384

DB 305 CAGACAGAAAGAACCGAGATCTGCGCTTCTGTAACGACTAGTGGCGCTCGAGA 364  
QY 385 AGTGACCTTAAGGCGGAAACACGCGCTCGAAGAGCATGCGACCCATTATGAGTT 444  
DB 365 AGTGACCTTAAGGCGGAAACACGCGCTCGAAGAGCATGCGACCCATTATGAGTT 424  
QY 445 CATCCAGACCTGGAGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 504  
DB 425 CATCCAGACCTGGAGAGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 484  
QY 505 GAGGAGCGAGATTAACAGCTCCAGCTCCGCTGCTTACACCGCAGATCGGAGGAGTT 564  
DB 485 GAGGAGCGAGATTAACAGCTCCAGCTCCGCTGCTTACACCGCAGATCGGAGGAGTT 544  
QY 565 ATAGTACCCGGGCTGGCTTACTACTAGTGTGATGATGATGATGATGATGATGATGAT 624  
DB 545 ATAGTACCCGGGCTGGCTTACTACTAGTGTGATGATGATGATGATGATGATGATGAT 604  
QY 625 GCTGTACTGTAAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 684  
DB 605 GCTGTACTGTAAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 664  
QY 685 GAATTCAGACCACTGCGGCGAGTTCCTGCGGCGCCAGCTCGGCTGCGAGTGTCT 744  
DB 665 GAATTCAGACCACTGCGGCGAGTTCCTGCGGCGCCAGCTCGGCTGCGAGTGTCT 724  
QY 745 GGGCTGTGGCGCTGCGGCGAGGCTTCCTGCGGAGTTCGACCTCCCTGGGCCAT 804  
DB 725 GGGCTGTGGCGCTGCGGCGAGGCTTCCTGCGGAGTTCGACCTCCCTGGGCCAT 784  
QY 805 CTCAGAGCTGCGGCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 864  
DB 785 CTCAGAGCTGCGGCGCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 844  
QY 865 GTCTCCCAAGTGTGCTGCGGCGAGGCTGCGGCGCTGCGAGAGTCTGCGGAGCCGCTC 924  
DB 845 GTCTCCCAAGTGTGCTGCGGCGAGGCTGCGGCGCTGCGAGAGTCTGCGGAGCCGCTC 904  
QY 925 CCTCTGCGGCGCTGCGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 984  
DB 905 CCTCTGCGGCGCTGCGGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 964  
QY 985 CTGGGCTGTGACGTTTTCATCCCATTAATTAATTAATTAATTAATTAATTAATTAAT 1044  
DB 965 CTGGGCTGTGACGTTTTCATCCCATTAATTAATTAATTAATTAATTAATTAATTAAT 1024  
QY 1045 AACTCCCGCAGCGCCACTCTCCACTCAGTCCCAATCCCTGAGCCCTTGTGAGGC 1104  
DB 1025 AACTCCCGCAGCGCCACTCTCCACTCAGTCCCAATCCCTGAGCCCTTGTGAGGC 1084  
QY 1105 CCCAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1164  
DB 1085 CCCAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1144  
QY 1165 TGTGGGCAAGAGTGGTTCAGAAAGCCACTTCAAGGCACTAAGAGGGGCTGGAGCT 1224  
DB 1145 TGTGGGCAAGAGTGGTTCAGAAAGCCACTTCAAGGCACTAAGAGGGGCTGGAGCT 1204  
QY 1225 GGCAGAGAACCAAGAGACTGAGGCTTAGGCAAGAGTCCCAATGTGAGGGGCGAGAA 1284  
DB 1205 GGCAGAGAACCAAGAGACTGAGGCTTAGGCAAGAGTCCCAATGTGAGGGGCGAGAA 1264  
QY 1285 CAAAGCAAGCTCTCTCTGAGATTCCTGCTGAGATTTTAAACAGATATATTTTAT 1344  
DB 1265 CAAAGCAAGCTCTCTCTGAGATTCCTGCTGAGATTTTAAACAGATATATTTTAT 1324  
QY 1345 TATTTATGACAAATGTTGATTAATGG 1373  
DB 1325 TATTTATGACAAATGTTGATTAATGG 1353

RESULT 3



AAA49717  
ID AAA49717 standard; cDNA, 1353 BP.  
XX  
AC AAA49717;  
XX  
DT 25-SEP-2000 (first entry)  
XX  
DE Human PRO207 cDNA clone DNA30879-1152.  
XX  
KM PRO207; human; antitumour; tumour; therapy; cytostatic;  
KM breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
KM uterine cancer; prostate cancer; lung cancer; bladder cancer;  
KM central nervous system cancer; melanoma; leukaemia; neoplasm; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 58..807  
FT sig\_peptide 58..177  
FT mat\_peptide 178..804  
FT /\*tag= a  
FT /\*tag= b  
FT /\*tag= c  
XX  
PN WO20037638-A2.  
XX  
PD 29-JUN-2000.  
XX  
PF 02-DEC-1999; 99WO-US28565.  
XX  
PR 22-DEC-1998; 98US-0113296.  
PR 08-MAR-1999; 99WO-US05028.  
PR 21-APR-1999; 99US-0130232.  
PR 28-APR-1999; 98US-0131445.  
PR 14-MAY-1999; 98US-0134287.  
PR 20-JUL-1999; 99US-0144758.  
PR 26-JUL-1999; 99US-0145698.  
PR 15-SEP-1999; 99WO-US21090.  
PR 15-SEP-1999; 99WO-US21547.  
XX  
PA (GENTECH ) GENENTECH INC.  
XX  
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;  
PI Napier MA, Pitti RM, Wood WI;  
XX  
DR WPI: 2000-442668/38.  
DR P-PSDB; AAY95338.  
XX  
PT Novel composition to inhibit neoplastic cell growth or for treating  
PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,  
PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO509 or  
PT PRO866  
XX  
PS Claim 20; Fig 3; 172pp; English.  
XX  
CC The present sequence is that of cDNA clone DNA30879-1152  
CC (ATCC 209358) encoding human PRO207 (see AAY95338), which shows  
CC homology to several members of the tumour necrosis factor family,  
CC especially human lymphotoxin (23.4%). The cDNA was identified in a  
CC foetal kidney cDNA library following identification of an expressed  
CC sequence tag with homology to human Apo-2 ligand. A claimed method  
CC for inhibiting the growth of a tumour cell comprises exposing the  
CC tumor cell to PRO179, PRO207, PRO320, PRO321, PRO224,  
CC PRO328, PRO301, PRO526, PRO362, PRO509 or PRO866 (see  
CC AAY95337-49), their agonists or chimeric polypeptides incorporating  
CC them. The tumour is especially a cancer selected from breast,  
CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and  
CC central nervous system cancer, melanoma and leukaemia. Nucleic  
CC acids encoding PRO179 etc. are used in the recombinant production  
CC of the antitumour polypeptides.  
XX  
SQ Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;

Query Match 96.2%; Score 1320.2; DB 21; Length 1353;  
Best Local Similarity 99.8%; Pred. No. 1.2e-271;  
Matches 1322; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 49 CGATCCCTCGGCTCCCGGGATGGGGGGCGGTGAGGCGACAGACCCGCCCCCATG 108  
DB 1 CGATCCCTCGGCTCCCGGGATGGGGGGCGGTGAGGCGACAGACCCGCCCCCATG 60  
QY 109 GCCGCCCTCGGAGCCAGCCAGGCGGAGGGGGCCCGGGGGGACCCGCCCTGCTG 168  
DB 61 GCCGCCCTCGGAGCCAGGCGGAGGGGGCCCGGGGGGACCCGCCCTGCTG 120  
QY 169 GTCCCGCTCGGCTGGGCGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGG 228  
DB 121 GTCCCGCTCGGCTGGGCGCTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGG 180  
QY 229 GTCACTTTGGGAGCCGGGCGATGCTGTCGCCCGCCAGAGCCCTGCCAGAGAGCTG 288  
DB 181 GTCACTTTGGGAGCCGGGCGATGCTGTCGCCCGCCAGAGCCCTGCCAGAGAGCTG 240  
QY 289 GCAGAGGAGGACAGGACCCCTGCGAATGATCCCGACAGAGAAAGCCAGATCT 348  
DB 241 GCAGAGGAGGACAGGACCCCTGCGAATGATCCCGACAGAGAAAGCCAGATCT 300  
QY 349 GCGCCTTCTGAACCGACTAGTGGGCTGCGAGAGTGACACTAAAGCCGAAACA 408  
DB 301 GCGCCTTCTGAACCGACTAGTGGGCTGCGAGAGTGACACTAAAGCCGAAACA 360  
QY 409 CGGGCTCGAAGAGCGATCGACGCCCATTTATGAAGTTTCATCCAGACTGGACAGCA 468  
DB 361 CGGGCTCGAAGAGCGATCGACGCCCATTTATGAAGTTTCATCCAGACTGGACAGCA 420  
QY 469 GCGGAGGAGGTGTGGAGCGGAGCAGTGAATGCTGGGAGAGAGCCAGATCAACAGCT 528  
DB 421 GCGGAGGAGGTGTGGAGCGGAGCAGTGAATGCTGGGAGAGAGCCAGATCAACAGCT 480  
QY 529 ACCCTCTGCTGCTACACCGCCAGATCGGGAGTTATGATACCGCGGCTGGGCTTAC 588  
DB 481 ACCCTCTGCTGCTACACCGCCAGATCGGGAGTTATGATACCGCGGCTGGGCTTAC 540  
QY 589 TACCTGTACTGTACAGTGATCTTGATGAGGGGAAAGCTGTCTACCTGAAGCTGAT 648  
DB 541 TACCTGTACTGTACAGTGATCTTGATGAGGGGAAAGCTGTCTACCTGAAGCTGAT 600  
QY 649 CTGGTGAATGTGTCTGGCCCTGCGGCGCTGGGAGGAAATTCAGCCATCGGCCAGT 708  
DB 601 CTGGTGAATGTGTCTGGCCCTGCGGCGCTGGGAGGAAATTCAGCCATCGGCCAGT 660  
QY 709 TCCCTCGGGCCCGACCTCGGCTGCGCAGGTGTGTGGGCTGTGGGCCCTGGCGCAGG 768  
DB 661 TCCCTCGGGCCCGACCTCGGCTGCGCAGGTGTGTGGGCTGTGGGCCCTGGCGCAGG 720  
QY 769 TCCCTCGGAGATCGGACCCCTCCCTGGGCCCATCTCAGGCTGCCCTTCTCAC 828  
DB 721 TCCCTCGGAGATCGGACCCCTCCCTGGGCCCATCTCAGGCTGCCCTTCTCAC 780  
QY 829 TACTCTCGAGCTTTCCAGATTCAGAGGAGGCGGCTGTCCCGACAGTGTCCAGGT 888  
DB 781 TACTCTCGAGCTTTCCAGATTCAGAGGAGGCGGCTGTCCCGACAGTGTCCAGGT 840  
QY 889 GCGGCGTCCCTCGACAGCTCTGTGGGACCGCGGCTCCCTGCGCCACCTCAGCGCT 948  
DB 841 GCGGCGTCCCTCGACAGCTCTGTGGGACCGCGGCTCCCTGCGCCACCTCAGCGCT 900  
QY 949 CTTCCTCGAGACTGCGCCCTCTCTGAGGCTCTGGGCTGTTCACGTGTTTCA 1008  
DB 901 CTTCCTCGAGACTGCGCCCTCTCTGAGGCTCTGGGCTGTTCACGTGTTTCA 960  
QY 1009 TCCCATTAATACAGTATTCCTCACTTTATTTACAACTCCCGACCGCCACTGTCA 1068  
DB 961 TCCCATTAATACAGTATTCCTCACTTTATTTACAACTCCCGACCGCCACTGTCA 1020  
QY 1069 COTCACTAGTCCCAATCCCGTGAACCTTTTGAAGGCCCGCAGTGAATCTGCCCTG 1128

Db 1021 CCTACTAGCTCCCAATCCCTGACCTTGGAGGCCCCCACTGATCTGCTCCCTG 1080  
 QY 1129 GCCACAGACCCCGAGGCGATGTGTCTACTGTCTGTGGGCAAGATGGTCCAGAG 1188  
 Db 1081 GCCACAGACCCCGAGGCGATGTGTCTACTGTCTGTGGGCAAGATGGTCCAGAG 1140  
 QY 1189 ACCCACTTACAGCCTAAGAGGGGCTGACCTGGCGGCAAGAACCAAGAGCTGGC 1248  
 Db 1141 ACCCACTTACAGCCTAAGAGGGGCTGACCTGGCGGCAAGAACCAAGAGCTGGC 1200  
 QY 1249 CTAGCCAGAGGTCCCAATGTGTAGGGGCGAGAAACAAGCAAGCTCTCCCTTGAGAA 1308  
 Db 1201 CTAGCCAGAGGTCCCAATGTGTAGGGGCGAGAAACAAGCAAGCTCTCCCTTGAGAA 1260  
 QY 1309 TTCCCTGTGATTTTAAACAGATATTATTATTATTATTATTATTATTATTATTATTATT 1368  
 Db 1261 TTCCCTGTGATTTTAAACAGATATTATTATTATTATTATTATTATTATTATTATTATT 1320  
 QY 1369 AATGG 1373  
 Db 1321 AATGG 1325

RESULT 5  
 ABR40255  
 ID ABR40255 standard; cDNA; 1353 BP.

XX ABR40255;

XX 15-JUL-2002 (first entry)

DE cDNA encoding human PRO207 polypeptide.

XX Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;

KM Leukemia; neuronal disorder; stromal disorder; blastocellic disorder;

KM Inflammatory disorder; immune disorder; angiogenic disorder;

KM gene therapy; cytostatic; neuroprotective; gene; ss.

OS Homo sapiens.

XX WO200153486-A1.

XX 26-JUL-2001.

XX 11-FEB-2000; 2000WO-US03565.

XX 08-MAR-1999; 99WO-US05028.

XX 11-MAR-1999; 99US-123972P.

XX 02-JUN-1999; 99WO-US12252.

XX 22-JUN-1999; 99US-140653P.

XX 20-JUL-1999; 99US-144758P.

XX 26-JUL-1999; 99US-145698P.

XX 28-JUL-1999; 99US-146222P.

XX 17-AUG-1999; 99US-149395P.

XX 31-AUG-1999; 99US-151689P.

XX 01-SEP-1999; 99WO-US20111.

XX 15-SEP-1999; 99WO-US21090.

XX 30-NOV-1999; 99WO-US28313.

XX 01-DEC-1999; 99WO-US28634.

XX 05-JAN-2000; 2000WO-US00219.

XX Thirty five nucleic acids encoding PRO polypeptides, useful for  
 PT treating benign or malignant tumours, leukemias and lymphoid  
 PT malignancies, inflammatory, angiogenic and immunologic disorders -  
 XX  
 PS Claim 50; Fig 3; 302pp: English.  
 XX  
 CC The present invention relates to the isolation of novel human PRO  
 CC polypeptides and the polynucleotide sequences encoding them. The  
 CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are  
 CC useful for treating benign or malignant tumours (e.g. renal, kidney,  
 CC bladder, breast, etc), leukemias and lymphoid malignancies, glandular,  
 CC disorders such as neuronal, glial, astrocytic, hypothalamic, glandular,  
 CC macrophagal, stromal and blastocellic disorders, inflammatory, immune  
 CC and angiogenic disorders. The polynucleotide sequences are also  
 CC useful in gene therapy. ABR40254-ABR40288 encode for the human PRO  
 CC polypeptides of the invention.  
 XX  
 SQ Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;

Query Match 96.2%; Score 1320.2; DB 24; Length 1353;  
 Best Local Similarity 99.8%; Pred. No. 1.2e-271;  
 Matches 1322; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 49 CGATCCCTCGGGTCCCGGATGGGGGCGGTGAGCGACAGCCCGCCCGCCATG 108  
 Db 1 CGATCCCTCGGGTCCCGGATGGGGGCGGTGAGCGACAGCCCGCCCGCCATG 60  
 QY 109 GCGCGCCGTGGAGCCAGAGCGGAGGGGCGCGGGGGAGCGCGGCGCCCGCTG 168  
 Db 61 GCGCGCCGTGGAGCCAGAGCGGAGGGGCGCGGGGGAGCGCGGCGCCCGCTG 120  
 QY 169 GTCCCGCTCGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGG 228  
 Db 121 GTCCCGCTCGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGGGCGTGG 180  
 QY 229 GTCACTTTGGGAGCGCGGCGATGCTCTCCGCCAGAGAGCTGCCAGAGAGCTG 288  
 Db 181 GTCACTTTGGGAGCGCGGCGATGCTCTCCGCCAGAGAGCTGCCAGAGAGCTG 240  
 QY 289 GCAGAGAGAGACAGAGACCGCGTGGAGTGAATCCAGACAGAGAGAGAGAGCT 348  
 Db 241 GCAGAGAGAGACAGAGACCGCGTGGAGTGAATCCAGACAGAGAGAGAGAGCT 300  
 QY 349 GCGCCTTCTCGAACCAGTATGTTGGCTCGCAGAGAGTCCACCTAAGCGCGAAACA 408  
 Db 301 GCGCCTTCTCGAACCAGTATGTTGGCTCGCAGAGAGTCCACCTAAGCGCGAAACA 360  
 QY 409 CGGCTCGAAGAGCGATCGACGCCATTATGAAGTTATCCAGAGCTGGACAGAGGA 468  
 Db 361 CGGCTCGAAGAGCGATCGACGCCATTATGAAGTTATCCAGAGCTGGACAGAGGA 420  
 QY 469 GCGCAGCAGGTGGAGCGGAGACAGTGAAGTGGAGGAGAGAGAGAGAGAGAGTCC 528  
 Db 421 GCGCAGCAGGTGGAGCGGAGACAGTGAAGTGGAGGAGAGAGAGAGAGAGAGTCC 480  
 QY 529 AGCCCTCGGCTAACAACCGCCAGATCGGGAGATTATAGTCAACCGCGGTGGCTTAC 588  
 Db 481 AGCCCTCGGCTAACAACCGCCAGATCGGGAGATTATAGTCAACCGCGGTGGCTTAC 540  
 QY 589 TACCTGACTGTCAAGTGCATTTGATGAGAGGAGAGAGCTCTCAACCTGAAGCTGACTTG 648  
 Db 541 TACCTGACTGTCAAGTGCATTTGATGAGAGGAGAGAGCTCTCAACCTGAAGCTGACTTG 600  
 QY 649 CTGGTGATGTGTGTGGGCTGGGCTGGGCTGGGAGAGATTCACGACCTAGCGGAGT 708  
 Db 601 CTGGTGATGTGTGTGGGCTGGGCTGGGCTGGGAGAGATTCACGACCTAGCGGAGT 660  
 QY 709 TCCTCGGGGCCAGCTCCGCTTCGCGAGAGTGTGGGCTGTGGCCCTGCGGCGAGG 768  
 Db 661 TCCTCGGGGCCAGCTCCGCTTCGCGAGAGTGTGGGCTGTGGCCCTGCGGCGAGG 720  
 QY 769 TCCTCCCTCGGAGTCCGACCCCTCCCTGGGCGCCATCTCAAGGCTGCCCTTCCTCACC 828

```

DB 721 TCTCTCGGATCGGACCCCTCCCTGGGCCAATCAAGCTGCCCTCCGAC 780
QY 829 TACTTCGAGCTTCCAGGTTCTACTGAGGGCCCTGCTGCCACAGTGTCCAGGCT 888
DB 781 TACTTCGAGCTTCCAGGTTCTACTGAGGGCCCTGCTGCCACAGTGTCCAGGCT 840
QY 889 GCGGGTCCCTCGAGAGTCTCTGGGACCGGGTCCCTGCTGCCACAGTGTCCAGGCT 948
DB 841 GCGGGTCCCTCGAGAGTCTCTGGGACCGGGTCCCTGCTGCCACAGTGTCCAGGCT 900
QY 949 CTTTGCCTCAGACTGCCCCCTCTCTAGAGGCTGCTGGGCTGTTCACGTGTTTCCA 1008
DB 901 CTTTGCCTCAGACTGCCCCCTCTCTAGAGGCTGCTGGGCTGTTCACGTGTTTCCA 960
QY 1009 TCCCATATAATACAGTATTCCTCACTCTTATTTACACTCCCGCCAGCTCTCCA 1068
DB 961 TCCCATATAATACAGTATTCCTCACTCTTATTTACACTCCCGCCAGCTCTCCA 1020
QY 1069 CCTCACTAGCTCCCAATCTGACCCCTTGAGGGCCCGCAGTGTGACTCCCGCTG 1128
DB 1021 CCTCACTAGCTCCCAATCTGACCCCTTGAGGGCCCGCAGTGTGACTCCCGCTG 1080
QY 1129 GCCACAGACCCCGAGGGCATTTGTTCACACTGACTCTGTGGCAAGATGGTCCAGAG 1188
DB 1081 GCCACAGACCCCGAGGGCATTTGTTCACACTGACTCTGTGGCAAGATGGTCCAGAG 1140
QY 1189 ACCCACTTCAGGCACTAAGAGGGGCTGACCTGGCGGACAGACCAAGAGACTGGGC 1248
DB 1141 ACCCACTTCAGGCACTAAGAGGGGCTGACCTGGCGGACAGACCAAGAGACTGGGC 1200
QY 1249 CTAGCCAGAGATTCCTCAATGTGAGGGGCGAGAAACAGACAGCTCCTCCTGAGAA 1308
DB 1201 CTAGCCAGAGATTCCTCAATGTGAGGGGCGAGAAACAGACAGCTCCTCCTGAGAA 1260
QY 1309 TTTCCCTGTGATTTTAAAAAGATATTTATTTATTTATTTATTTATTTATTTATTT 1368
DB 1261 TTTCCCTGTGATTTTAAAAAGATATTTATTTATTTATTTATTTATTTATTTATTTATTT 1320
QY 1369 AATGG 1373
DB 1321 AATGG 1325

```

RESULT 6  
AAV47613  
ID AAV47613 standard; cDNA; 1236 BP.

```

AC AAV47613;
XX
DT 27-OCT-1998 (first entry)
XX
DE TNF related endothelium proliferative agent gene.
XX
KW ss: TNF; endothelium proliferative agent; TREPA; wound healing; cancer;
XX tissue grafting; vascularisation; apoptosis; autoimmune; birth control.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..750
FT /tag= a
FT /product= "TREPA"
XX
PN W09835061-A2.
XX
PD 13-AUG-1998.
XX
PF 12-FEB-1998; 98WO-US02859.
XX
PR 10-FEB-1998; 98US-0021706.
XX 12-FEB-1997; 97US-0798692.
XX

```

```

PA (ABRO ) ABBOTT LAB.
PI Wiley SR;
XX WPI: 1998-447255/38.
XX P-PSDI: AAM29745.
XX
PT Detecting nucleic acid encoding TREPA - useful for diagnosis and
PT treatment of autoimmune disease, tumours and inflammation
XX
PS Claim 11: Page 123-4; 142pp; English.
XX
XX The TNF-related endothelium proliferative agent (TREPA), or its
XX activators or agonists, are used to treat a deficit of TREPA, e.g. to
XX promote wound healing or tissue grafting, by promoting vascularisation,
XX also to induce apoptosis for treating cancer and eliminating autoreactive
XX T cells, as an adjunct to cancer chemotherapy or antiviral treatment.
XX TREPA peptides can also be used to target cytotoxic agents or for
XX affinity isolation of the corresponding receptor, the nucleic acid for
XX which can be used to transform tumour cells to render them more
XX responsive to TREPA and to screen for TREPA mimics.
XX Ribozymes, antisense RNA, antibodies or peptides, are used to treat
XX (vascularisation), inflammation or a wide range of autoimmune conditions,
XX conditions involving abnormal stimulation of epithelial cells (e.g.
XX atherosclerosis), for birth control (inhibiting ovulation and placental
XX formation) or other angiogenic conditions (e.g. ulcers).
XX
XX Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;
XX
Query Match 89.3%; Score 1226.4; DB 19; Length 1236;
Best Local Similarity 99.5%; Pred. No. 1-Le-251;
Matches 1230; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 106 ATGGCCGCCCGCTCGAGCCAGAGCGGAGGGGCGCCGGGGAGCCGGGACCCGCTG 165
DB 1 ATGGCCGCCCGCTCGAGCCAGAGCGGAGGGGCGCCGGGGAGCCGGGACCCGCTG 60
QY 166 CTGGTCCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 225
DB 61 CTGGTCCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 120
QY 226 GTGGTCAGTTTGGGAGCGGGCATGCTGTCCGCCAGAGACTGCGCCAGAGAGGCTG 285
DB 121 GTGGTCAGTTTGGGAGCGGGCATGCTGTCCGCCAGAGACTGCGCCAGAGAGGCTG 180
QY 286 GTGGCAGAGAGAGACAGACCCGTCGGAATGAATCCCGACAGAAAGAGAGAT 345
DB 181 GTGGCAGAGAGAGACAGACCCGTCGGAATGAATCCCGACAGAAAGAGAGAT 240
QY 346 CCTGGCCCTTCTTCTGACCGAGTATTTGGGCTGCGAGAAATGACCTTAAGGCCGAGAA 405
DB 241 CCTGGCCCTTCTTCTGACCGAGTATTTGGGCTGCGAGAAATGACCTTAAGGCCGAGAA 300
QY 406 ACAGGGGCTCGAAGAGCGATGCGACGCCATTTGAAGTTTCAATCCAGAGCTGAGAGAGC 465
DB 301 ACAGGGGCTCGAAGAGCGATGCGACGCCATTTGAAGTTTCAATCCAGAGCTGAGAGAGC 360
QY 466 GGAGCGCAGAGCGAGTGTGACGAGAGTGTGAGTGGCTGGGAGAGCCAGAAATCAACAGC 525
DB 361 GGAGCGCAGAGCGAGTGTGACGAGAGTGTGAGTGGCTGGGAGAGCCAGAAATCAACAGC 420
QY 526 TCCAGCCCTCTGCGCTACACCGCCAGATCGGGAGTTTATATGTAACCCGGGCTGGGCTC 585
DB 421 TCCAGCCCTCTGCGCTACACCGCCAGATCGGGAGTTTATATGTAACCCGGGCTGGGCTC 480
QY 586 TACTACCTGTACTGTCAAGTGCATTTGATGAGGGGAGAGGCTGTACTGTGAAGTGGAC 645
DB 481 TACTACCTGTACTGTCAAGTGCATTTGATGAGGGGAGAGGCTGTACTGTGAAGTGGAC 540
QY 646 TTGCTGTGAGATGTGTCTGGCCCTGCGCTGCTGAGAGAAATTCACAGCCACTGCGGCC 705
DB 541 TTGCTGTGAGATGTGTCTGGCCCTGCGCTGCTGAGAGAAATTCACAGCCACTGCGGCC 600

```

QY 706 AGTTCCTCGGGCCCAAGCTCCGCTCTGCGAGGTGTCGGCTGTGGCCCTCGCGCA 765  
 |||||  
 Db 601 AGTTCCTCGGGCCCAAGCTCCGCTCTGCGAGGTGTCGGCTGTGGCCCTCGCGCA 660  
 QY 766 GGGTCTTCCTCGGGATTCGACCTCCCTGGGCGCATCTCAAGGTGCCCCCTCTCTC 825  
 |||||  
 Db 661 GGGTCTTCCTCGGGATTCGACCTCCCTGGGCGCATCTCAAGGTGCCCCCTCTCTC 720  
 QY 826 ACCTACTCGGACTCTTCGACAGTTCAGTGAAGGGCCCTGCTCCCACTGCTGCGAG 885  
 |||||  
 Db 721 ACCTACTCGGACTCTTCGACAGTTCAGTGAAGGGCCCTGCTCCCACTGCTGCGAG 780  
 QY 886 GCTGCGGCTCCCTCGACAGCTCTGCGGACCCGGTCCCTCTGCGCCACCTCAAGCC 945  
 |||||  
 Db 781 GCTGCGGCTCCCTCGACAGCTCTGCGGACCCGGTCCCTCTGCGCCACCTCAAGCC 840  
 QY 946 GCTCTTTCGTCAGACCTCCGCTCCCTCTGAGGCTGCGGCGCTGTCTACGTGTTT 1005  
 |||||  
 Db 841 GCTCTTTCGTCAGACCTCCGCTCCCTCTGAGGCTGCGGCGCTGTCTACGTGTTT 900  
 QY 1006 CCATCCACATTAATACAGTATTCCTCACTCTTATCTTACACTCCCGACCGCCACTCT 1065  
 |||||  
 Db 901 CCATCCACATTAATACAGTATTCCTCACTCTTATCTTACACTCCCGACCGCCACTCT 960  
 QY 1066 CCACCTCACTAGCTCCCAATCCCTGACCCCTTGAAGGCGCCAGTATCTGACTCCCC 1125  
 |||||  
 Db 961 CCACCTCACTAGCTCCCAATCCCTGACCCCTTGAAGGCGCCAGTATCTGACTCCCC 1020  
 QY 1126 CTGCGCACAGACCCCGAGGGGATGTGTCTACCTGCTGCTGCGCAAGATGGGTCCAG 1185  
 |||||  
 Db 1021 CTGCGCACAGACCCCGAGGGGATGTGTCTACCTGCTGCTGCGCAAGATGGGTCCAG 1080  
 QY 1186 AAGACCCCACTTCAAGGACTAAGAGGGGCTGACCTGCGCGCAGAAAGCCAAAGAGACTG 1245  
 |||||  
 Db 1081 AAGACCCCACTTCAAGGACTAAGAGGGGCTGACCTGCGCGCAGAAAGCCAAAGAGACTG 1140  
 QY 1246 GGCCTAGGCGGAGATTCCTCAATGTGAGGGGCGAGAAACAGACAGCTCTCTCTTGA 1305  
 |||||  
 Db 1141 GGCCTAGGCGGAGATTCCTCAATGTGAGGGGCGAGAAACAGACAGCTCTCTCTTGA 1200  
 QY 1306 GAATTCCTGTGATTTTAAACAGATATATTTT 1341  
 |||||  
 Db 1201 GAATTCCTGTGATTTTAAACAGATATATTTT 1236  
 RESULT 7  
 AAD04350  
 ID AAD04350 standard; cDNA; 1236 BP.  
 XX  
 AC AAD04350;  
 XX  
 DT 04-JUL-2001 (first entry)  
 XX  
 DE Human TREPA (TNF related endothelium proliferative agent) cDNA.  
 XX  
 KW Human: tumour necrosis factor; TNF; angiogenesis; wound healing;  
 KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;  
 KW grafting; vlnlnerary; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..750  
 FT CDS /\*tag- a  
 FT /product= "Human TREPA (TNF related endothelium  
 FT proliferative agent)"  
 XX  
 PN US6207642-B1.  
 XX  
 PD 27-MAR-2001.  
 XX  
 PF 26-JUN-1998; 98US-0105343.

XX  
 PR 12-FEB-1997; 97US-0798692.  
 PR 10-FEB-1998; 98US-0021706.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 PA  
 PI Willey SR;  
 PI  
 DR WPI; 2001-280760/29.  
 DR P-PSDB; AAE00891.  
 XX  
 PT Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 PT  
 PS Example 2; Column 73-74; 53pp; English.  
 XX  
 CC The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts.  
 CC The present sequence is a cDNA clone ID #69050 encoding human TREPA.  
 CC  
 SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;  
 Query Match 89.3%; Score 1226.4; DB 22; Length 1236;  
 Best Local Similarity 99.58; Pred. No. 1..le-251;  
 Matches 1230; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
 QY 106 ATGCGCCCGCTCGAGACCAAGAGGGGGCGCGGGGAGCGCGGACCGCCCTG 165  
 |||||  
 Db 1 ATGCGCCCGCTCGAGACCAAGAGGGGGCGCGGGGAGCGCGGACCGCCCTG 60  
 QY 166 CTGCTCCCGCTCGGCTGGGCGCTGGGCGCTGGCGCTGCTGCTGCTGCGCC 225  
 |||||  
 Db 61 CTGCTCCCGCTCGGCTGGGCGCTGGGCGCTGGCGCTGCGCTGCTGCTGCGCC 120  
 QY 226 GTGCTCAGTTTGGGAGACCGCGGCTGCTGCTGCGCCAGAGAGCTGCGCCAGAGACTG 265  
 |||||  
 Db 121 GTGCTCAGTTTGGGAGACCGCGGCTGCTGCTGCGCCAGAGAGCTGCGCCAGAGACTG 180  
 QY 286 GTGCGAAGAGAGACAGACAGACCGCTCGGAATGATCCAGACAGAAAGCCAGAT 345  
 |||||  
 Db 181 GTGCGAAGAGAGACAGACAGACCGCTCGGAATGATCCAGACAGAAAGCCAGAT 240  
 QY 346 CTGCGCCCTTCTGAACCGACTAGTTCGCTGCGCAGAAAGTGCACCTAAAGCCGGA 405  
 |||||  
 Db 241 CTGCGCCCTTCTGAACCGACTAGTTCGCTGCGCAGAAAGTGCACCTAAAGCCGGA 300  
 QY 406 ACAGGGGCTCGAAGAGAGGATCGGAGCCCATTAAGATTCACAGACTGAGAGAG 465  
 |||||  
 Db 301 ACAGGGGCTCGAAGAGAGGATCGGAGCCCATTAAGATTCACAGACTGAGAGAG 360  
 QY 466 GAGAGCGAGCAGAGTGTGAGAGGAGACAGTAGTGTGAGAGAGAGCAGAAATCAACAGC 525  
 |||||  
 Db 361 GAGAGCGAGCAGAGTGTGAGAGGAGACAGTAGTGTGAGAGAGAGCAGAAATCAACAGC 420  
 QY 526 TCCAGCCCTTGGGCTTCAACACCGCCAGATCGGGAGTTTATAGTCAACCGGCTGGGCTC 585  
 |||||  
 Db 421 TCCAGCCCTTGGGCTTCAACACCGCCAGATCGGGAGTTTATAGTCAACCGGCTGGGCTC 480  
 QY 586 TACTACTGTAACCTCAAGTGTGACATTTGATGAGAGGAGAGGCTGTACTCTGAAGCTGAGAC 645  
 |||||  
 Db 481 TACTACTGTAACCTCAAGTGTGACATTTGATGAGAGGAGAGGCTGTACTCTGAAGCTGAGAC 540  
 QY 646 TTGCTGTGATGTGTGTGCTGCGCTGCTGCTGAGAGAAATTCAGCACTGCGGCG 705  
 |||||  
 Db 541 TTGCTGTGATGTGTGTGCTGCGCTGCTGCTGAGAGAAATTCAGCACTGCGGCG 600



```

OY 706 ACTTCCTCGGGGCCCCAGCTCCGCTCTGCGCAGTGTCTGGGCTGTGGCCCTGGGGCCA 765
    |||||||
DB 601 AGTTCCTCGGGGCCCCAGCTCCGCTCTGCGCAGTGTCTGGGCTGTGGCCCTGGGGCCA 660
OY 766 GGGTCTCTCCCTGGGATCGGACACCTCTCCCTGGGCCCCATCTCAGAGCTGCCCTTCTCT 825
    |||||||
DB 661 GGGTCTCTCCCTGGGATCGGACACCTCTCCCTGGGCCCCATCTCAGAGCTGCCCTTCTCT 720
OY 826 ACCTACTCTGGAGCTCTTCCAGGTCTACAGAGGGGGCCCTGTGCTCCCGACAGTGTCCAG 885
    |||||||
DB 721 ACCTACTCTGGAGCTCTTCCAGGTCTACAGAGGGGGCCCTGTGCTCCCGACAGTGTCCAG 780
OY 886 GCTGCGGGCTCCCTCGACAGCTCTCTGGGACCGGGTCCCTCTGCCCCACCTCCAGCC 945
    |||||||
DB 781 GCTGCGGGCTCCCTCGACAGCTCTCTGGGACCGGGTCCCTCTGCCCCACCTCCAGCC 840
OY 946 GCTCTTTGCTCCAGACCTGCCCCCTCTCTAGAGGCTGCTGGGCTCTTCCAGTGTCTT 1005
    |||||||
DB 841 GCTCTTTGCTCCAGACCTGCCCCCTCTCTAGAGGCTGCTGGGCTCTTCCAGTGTCTT 900
OY 1006 CCATCCCACTAAATACAGTATCTCCACTCTTATCTTACAACTCCCGCCACCCACTCT 1065
    |||||||
DB 901 CCATCCCACTAAATACAGTATCTCCACTCTTATCTTACAACTCCCGCCACCCACTCT 960
OY 1066 CCACCTCACTAGCTCCCAATCTCCAGCTTGTGAGGCCCCCAGTATCTGACTCTCCCC 1125
    |||||||
DB 961 CCACCTCACTAGCTCCCAATCTCCAGCTTGTGAGGCCCCCAGTATCTGACTCTCCCC 1020
OY 1126 CTGGCCACAGACCCCGAGGCACTGTGTCTACTCTGTGTGGCCAGAGATGGTCCAG 1185
    |||||||
DB 1021 CTGGCCACAGACCCCGAGGCACTGTGTCTACTCTGTGTGGCCAGAGATGGTCCAG 1080
OY 1186 AAGACCCCACTGAGGCACTAGAGGGGCTGAGCTGGGCGGAGGAAGCAAGAGACTG 1245
    |||||||
DB 1081 AAGACCCCACTGAGGCACTAGAGGGGCTGAGCTGGGCGGAGGAAGCAAGAGACTG 1140
OY 1246 GGCCCTAGGCGAGGATCTCCCAATGTGAGGGGCGAGAAACAAGACAAGCTCTCCCTTGA 1305
    |||||||
DB 1141 GGCCCTAGGCGAGGATCTCCCAATGTGAGGGGCGAGAAACAAGACAAGCTCTCCCTTGA 1200
OY 1306 GAATTCCTCTGTGGATTTTAAACAGATATATTTT 1341
    |||||||
DB 1201 GAATTCCTCTGTGGATTTTAAACAGATATATTTT 1236

```

## RESULT 8

AAAX23424  
ID AAAX23424 standard; DNA; 1030 BP.

AAAX23424;  
18-JUN-1999 (first entry)

Human TNRL3 DNA.

Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;  
developmental abnormality; gestational abnormality; prostate cancer;  
APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
apoptosis; human; ss.

Homo sapiens.

Key Location/Qualifiers  
1..627  
/\*tag= a  
/product= "TNRL3"

W09911791-A2.

11-MAR-1999.

04-SEP-1998; 98WO-US18393.

```

XX 05-SEP-1997; 97US-0924634.
XX (UNIV ) UNIV WASHINGTON.
XX Chaudhary PM;
XX WPI: 1999-205191/17.
XX P-PSDB: AAM93590.
XX New Tumor Necrosis Factor family receptor polypeptides and ligands -
XX useful for diagnosis and treatment of prostate cancer and
XX developmental or gestational abnormalities
XX Example VII; Fig 13a. 156pp; English.
XX This invention describes isolated Tumor Necrosis Factor (TNF) family
XX receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
XX fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
XX their active fragments. APO4 is useful for diagnosing prostate cancer
XX by determining levels of APO4 in an individual. Prostate cancer can also
XX be treated using APO4 selective binding agents linked to a therapeutic
XX moiety. APO4 polypeptides are also useful for identifying selective
XX binding agents, useful in diagnosis/treatment of disease by binding of
XX agents to the polypeptide/active fragment which is extracellular, or
XX expressed on the cell surface. The binding is preferably performed in
XX vivo. APO4 polypeptides/active fragments are also useful for screening
XX for agonists and antagonists by binding and observing the change in APO4
XX activity. Effective pharmacological agents useful in diagnosis or
XX treatment of disease are also identified using APO4 polypeptides/active
XX fragments and APO4 signal transducer molecules that specifically interact
XX with a cytoplasmic domain of APO4 and detecting a change in level of APO4
XX activity. The method is performed in vivo or in vitro. APO polypeptides
XX are all useful as immunogens for preparing antibodies. APO4 is also
XX useful for diagnosis/treatment of developmental or gestational
XX abnormalities. APO8 was transfected to human breast carcinoma cell line
XX MCF-7, and induced apoptosis.

```

Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;

Query Match 60.7%; Score 833.4; DB 20; Length 1030;  
Best Local Similarity 99.9%; Pred. No. 4.8e-168;  
Matches 834; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 229 CTCAGTTTGGGAGCGCGGACATCGCTGTCGCCCGCAGAGACCTGCCAGAGAGCTGGTG 288
    |||||||
DB 1 GTCAGTTTGGGAGCGCGGACATCGCTGTCGCCCGCAGAGACCTGCCAGAGAGCTGGTG 60
OY 289 GCAGAGGAGACACGAGCCCGTCGGAAGTGAATCCCGCAGACAGAGAAAGCCAGATCCT 348
    |||||||
DB 61 GCAGAGGAGACACGAGCCCGTCGGAAGTGAATCCCGCAGACAGAGAAAGCCAGATCCT 120
OY 349 GCGCTTTCCTGAAACGACTAGTGGGCTCGCAGAGTGCACCTAAAGCCGGAANAACA 408
    |||||||
DB 121 GCGCTTTCCTGAAACGACTAGTGGGCTCGCAGAGTGCACCTAAAGCCGGAANAACA 180
OY 409 CGGGCTCGAAGAGCATGGCAGCCCATATATGAAGTTTCAACGACCTGAGACAGACGA 468
    |||||||
DB 181 CGGGCTCGAAGAGCATGGCAGCCCATATATGAAGTTTCAACGACCTGAGACAGACGA 240
OY 469 GCGCAGGAGAGTGTGAGCGGAGCAGTGAATGGCTGGGAGGAAGCCAGATCAACAGCTCC 528
    |||||||
DB 241 GCGCAGGAGAGTGTGAGCGGAGCAGTGAATGGCTGGGAGGAAGCCAGATCAACAGCTCC 300
OY 529 AGCCCTCTGCGCTACACCCCGCAGATCGGGAGTTATAGTACACCCGGGCTGGCTTAC 588
    |||||||
DB 301 AGCCCTCTGCGCTACACCCCGCAGATCGGGAGTTATAGTACACCCGGGCTGGCTTAC 360
OY 589 TACCTGTACTGTACAGTGCATTTTATATAGAGGGAAGGCTGTCTACTCTAAGCTGACTTG 648
    |||||||
DB 361 TACCTGTACTGTACAGTGCATTTTATATAGAGGGAAGGCTGTCTACTCTAAGCTGACTTG 420
OY 649 CTGCTGATGTGTGCTGGCCCTGGCTGCTGAGAGAAATCTCAGCCATGCGGCCAGT 708

```

```

Db 421 CTGGTGGATGGTGTCTCTGGCCCTGGCCCTGCTGAGAGATTTCTGAGCCACTGGCCACT 480
Qy 709 TCCCTGGGGCCCGACCTCCGCTCTGCGAGGTGTCTGGAGCTGTGGCCCTGGCCAGAGG 768
Db 481 TCCCTGGGGCCCGACCTCCGCTCTGCGAGGTGTCTGGAGCTGTGGCCCTGGCCAGAGG 540
Qy 769 TCCCTGGGGCCCGAGTCCCTCCGCTGGCCCTGCTGAGAGTGTCTGGAGCTGTGGCCCT 828
Db 541 TCCCTGGGGCCCGAGTCCCTCCGCTGGCCCTGCTGAGAGTGTCTGGAGCTGTGGCCCT 600
Qy 829 TACTTGGAGCTCTTCCAGAGTCTGAGAGGCGCTGCTGCTCCAGAGTGTGCGAGGCT 888
Db 601 TACTTGGAGCTCTTCCAGAGTCTGAGAGGCGCTGCTGCTCCAGAGTGTGCGAGGCT 660
Qy 889 GCGGCTCCCTCGACAGCTCTGCGAGGACCGGCTCCCTCTGCGCCAGCTCCAGCCGCT 948
Db 661 GCGGCTCCCTCGACAGCTCTGCGAGGACCGGCTCCCTCTGCGCCAGCTCCAGCCGCT 720
Qy 949 CTTTGTCCAGAGCTCCGCTCCGCTGAGAGGCTGCGCGGCTGCTGAGAGTGTGCGAGGCT 1008
Db 721 CTTTGTCCAGAGCTCCGCTCCGCTGAGAGGCTGCGCGGCTGCTGAGAGTGTGCGAGGCT 780
Qy 1009 TCCGACATTAATACAGTATTCCTCTTATCTTACAACTCCCGCCAGCCGCT 1063
Db 781 TCCGACATTAATACAGTATTCCTCTTATCTTACAACTCCCGCCAGCCGCT 835

```

## RESULT 9

AA03964  
ID AA03964 standard; DNA: 898 BP.

AC AA03964:

DT 26-SEP-2001 (first entry)

DE Expression vector pDC409-L2-TWEAK fusion protein-encoding DNA.

XX TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;  
 KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 KW rubeosis; uveitis; macular degeneration; arthritis; rheumatism; ds;  
 KW corneal graft neovascularisation; psoriasis; metastatic condition;  
 KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
 KW preneoplastic condition; myocardial angiogenesis; wound granulation;  
 KW scleroderma; vascular adhesion; telangiectasia; ischemia; human;  
 KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 KW peripheral atherosclerosis; pDC409-L2-TWEAK; TWEAK receptor; TWEAKR;  
 KW fusion protein.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

PH 52..873

FT /\*tag- a "Fusion protein comprising a growth hormone

FT leader, a leucine zipper multimerisation

FT domain, and human TWEAK extracellular

domain"

XX MO200145730-A2.

XX PD 28-JUN-2001.

XX PF 19-DEC-2000; 2000MO-US34755.

XX PR 20-DEC-1999; 99US-0172878.

XX PR 10-MAY-2000; 2000US-0203347.

XX (IMMUNEX CORP.

XX PA Willey SR;

XX PI

XX WP1: 2001-417975/44.  
 DR P-PSDB: AAU03499.  
 XX  
 PT Modulating angiogenesis in a mammal for treating diseases mediated by  
 PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  
 PT peripheral tissue, by administering antagonist or agonist of TWEAK  
 PT receptor  
 PS  
 XX Example 1: Page 39-40; 46pp; English.

The sequence represents a DNA from the expression vector  
 CC pDC409-L2-TWEAK, which encodes a fusion protein comprising a growth  
 CC hormone leader, a leucine zipper multimerisation domain, and the  
 CC extracellular domain of human TWEAK. The fusion protein was used in  
 CC the isolation of human TWEAK receptor (TWEAKR)-expressing clones  
 CC from a COS cell human cDNA library. The TWEAK protein is a  
 CC member of the tumour necrosis factor (TNF) family and induces  
 CC angiogenesis. TWEAKR may therefore be used to screen for and develop  
 CC TWEAKR agonists and antagonists for the modulation of angiogenesis, to be  
 CC used in the treatment and diagnosis of human disease. The disorders  
 CC mediated by angiogenesis include ocular disorders characterised by ocular  
 CC neovascularisation such as diabetic retinopathy, neovascular glaucoma,  
 CC retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,  
 CC rubeosis, uveitis, macular degeneration and corneal graft  
 CC neovascularisation, and inflammatory diseases such as arthritis,  
 CC rheumatism and psoriasis. Other treatable diseases include malignant and  
 CC metastatic conditions such as sarcomas and carcinomas, benign tumours and  
 CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,  
 CC scleroderma, vascular adhesions, atherosclerotic plaque  
 CC neovascularisation, telangiectasia, wound granulation, coronary  
 CC atherosclerosis, peripheral atherosclerosis and ischemia.  
 CC  
 XX Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other;

Query Match 45.88; Score 629.2; DB 22; Length 898;  
 Best Local Similarity 99.58; Pred. No. 1,4e-124;  
 Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

Qy 232 AGTTGGGAGCCCGGACATCCCTGCTGCGCCAGAGAGCTGCGCCAGAGAGCTGTGGCA 291
Db 250 AGTTGGGAGCCCGGACATCCCTGCTGCGCCAGAGAGCTGCGCCAGAGAGCTGTGGCA 309
Qy 292 GAGGAGGACGAGCCGCTGCGGAACTGATCCCAAGAGAGAGAGAGAGAGAGAGAGAG 351
Db 310 GAGGAGGACGAGCCGCTGCGGAACTGATCCCAAGAGAGAGAGAGAGAGAGAGAGAG 369
Qy 352 CCTTCCGAAACGAGTATGTTGCGCTGCGGAGAGTGCACCTAAGCCGGAAGAACGCG 411
Db 370 CCTTCCGAAACGAGTATGTTGCGCTGCGGAGAGTGCACCTAAGCCGGAAGAACGCG 429
Qy 412 GCTGGAAGAGGATCGAGCCATTAATGAATTATCAGACCTGAGAGAGAGAGAGAGAG 471
Db 430 GCTGGAAGAGGATCGAGCCATTAATGAATTATCAGACCTGAGAGAGAGAGAGAGAG 489
Qy 472 CAGCAGAGTGTGAGCGGAGAGAGTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGAG 531
Db 490 CAGCAGAGTGTGAGCGGAGAGAGTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGAG 549
Qy 532 CCTGCGGCTAACAACGCGCAGATCGGGAGTTTATATCAACCGGGCTGGGCTCTACTAC 591
Db 550 CCTGCGGCTAACAACGCGCAGATCGGGAGTTTATATCAACCGGGCTGGGCTCTACTAC 609
Qy 592 CTGACTGTGAGTGTGACTTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 651
Db 610 CTGACTGTGAGTGTGACTTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 669
Qy 652 GTGAGTGTGTGCTGCGCTGCGCTGCGTGGAGAGATTTTCAGACCTGCGGCAAGTTC 711
Db 670 GTGAGTGTGTGCTGCGCTGCGCTGCGTGGAGAGATTTTCAGACCTGCGGCAAGTTC 729
Qy 712 CTGCGGCGGAGCTCCGCTCTGCGAGGTGTGTGGCTGTGGCCCTGCGGCGAGGCTCC 771

```







Db 181 TCCGTGATTTT 195

RESULT 14  
AA53491/c  
ID AA53491 standard; DNA; 114955 BP.  
XX  
AC AA53491;  
XX  
DT 05-JUL-1999 (first entry)  
XX  
DE Human adenosine A1 receptor antisense oligonucleotide fragment.  
XX  
KW Antisense oligonucleotide; multiple target; antisense treatment;  
KW impaired respiration; inflammation; lung disease;  
KW pulmonary vasoconstriction; inflammation; allergic rhinitis;  
KW acute asthma; allergy; asthma; impaired respiration;  
KW respiratory distress syndrome; pain; cystic fibrosis;  
KW pulmonary hypertension; pulmonary vasoconstriction; emphysema;  
KW chronic obstructive pulmonary disease; leukemia; lymphoma; carcinoma;  
KW colon cancer; breast cancer; lung cancer; pancreatic cancer;  
KW hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;  
KW prostate cancer; ss.  
XX  
OS Synthetic.  
XX  
PN WO911886-A1.  
XX  
PD 25-MAR-1999.  
XX  
PF 17-SEP-1998; 98WO-US19419.  
XX  
PR 09-JUN-1998; 98US-0093972.  
PR 17-SEP-1997; 97US-0059160.  
XX  
PA (UYEC-) UNIV EAST CAROLINA.  
XX  
PI Myce JM;  
XX  
DR WPI; 1999-229400/19.  
XX  
PT New antisense oligonucleotides used in treatment of, e.g. pulmonary  
PT vasoconstriction  
XX  
PS Disclosure; Page 37; 120pp; English.  
XX  
CC The specification describes antisense oligonucleotides (AA52869-X55271)  
CC directed against at least 2 mRNAs selected from target genes, coding and  
CC non-coding regions of RNAs corresponding to target genes, gene  
CC initiation codons, genomic flanking regions, intron-exon borders, the  
CC 5'-end, the 3'-end and the juxta-section between coding and non-coding  
CC regions and all segments of RNAs encoding proteins associated with one  
CC or more diseases, conditions or mixtures. The antisense oligonucleotides  
CC may be derived from sequences AA5272-74. These multiple target  
CC oligonucleotides (specifically AA5180-271) can be used for the  
CC antisense treatment of diseases and conditions. Typical diseases and  
CC conditions are those associated with impaired respiration and  
CC inflammation, including lung diseases, pulmonary vasoconstriction,  
CC inflammation, allergic rhinitis, acute asthma, cystic fibrosis,  
CC respiration, respiratory distress syndrome, pain, lymphoma, impaired  
CC obstructive pulmonary disease (COPD), and cancers such as leukemias,  
CC lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer,  
CC hepatic metastases, as well as all types of cancers which may metastasize  
CC or have metastasized to the lungs, including breast and prostate cancer.  
XX  
SQ Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;

Query Match 5.18; Score 70.4; DB 20; Length 114955;  
Best Local Similarity 32.78; Pred. No. 3.6e-05;  
Matches 182; Conservative 58; Mismatches 316; Indels 0; Gaps 0;

OY 26 CCGCCCGCGGCTCCCTCCCGATCTTCGGGTCCCGGATGGGGGCGGTGAGGC 85  
||| ||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 105272 CCGCCCGGCGCGCCNNNNNNNSCGCCCGCGCGGCGCCNNNNNNNSCGCCCGC 105213  
OY 86 AGGCACAGCCCCCCCCCAATGCGCCCGCTCGAGACAGAGCGGAGCGGCCCGC 145  
||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 105212 CGGCGCGCGCCNNNNNNNSCGCCCGCGCGGCGCCNNNNNNNSCGCCCGC 105153  
OY 146 GGGAGCGGGGACCGCCCTGTGCTGCTCGGCTCGGCTGGGCTGGGCTGGGCT 205  
||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 105152 GGGCGCGCGCCNNNNNNNSCGCCCGCGCGGCGCCNNNNNNNSCGCCCGC 105093  
OY 206 GCTTCGCTCTCTGCTGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 265  
||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 105092 GCGCGCGCGCGCCNNNNNNNSCGCCCGCGCGGCGCCNNNNNNNSCGCCCGC 105033  
OY 266 AGCCGCGCGAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 325  
||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 105032 CCG 104973  
OY 326 AGACAGAGAAAGAGGATCTGCGCTTCTGTAACGACTAGTGTGCTGCAAGAA 385  
||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 104972 CCGGNNNNNNNSCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104913  
OY 386 GTGACCTTAAAGCGCGGAAACACBGGCTCGAAGAGCATGCAAGCCATTATGAATT 445  
||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 104912 CGGCG 104853  
OY 446 ATCCAGACCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 505  
||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 104852 VGGCCVGGGNNNNNSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104793  
OY 506 AGAAGCCAGATCAATCAAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 565  
||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 104792 VGGCCVGGGNNNNNSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 104733  
OY 566 TAGTCACCCCGGCGTGG 581  
||| ||| ||| : : : ||| ||| ||| : : : ||| |||  
Db 104732 NSCGCGCGCGCGCGCG 104717

RESULT 15  
AA217263  
ID AA217263 standard; cDNA; 1337 BP.  
XX  
AC AA217263;  
XX  
DT 12-OCT-1999 (first entry)  
XX  
DE Human gene expression product cDNA sequence SEQ ID NO:4735.  
KW Human; gene; gene expression product; diagnosis; therapy; probe;  
KW detection; mapping; tissue typing; profiling; forensic; cancer;  
KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO938972-A2.  
XX  
PD 05-AUG-1999.  
XX  
PF 28-JAN-1999; 99WO-US01619.  
XX  
PR 03-APR-1998; 98US-0080666.  
PR 28-JAN-1998; 98US-0072910.  
PR 24-FEB-1998; 98US-0075954.  
PR 31-MAR-1998; 98US-0080114.  
PR 03-APR-1998; 98US-0080515.  
XX  
PA (CHIR) CHIRON CORP.  
PA (HYSE-) HYSEQ INC.  
XX  
PI Ctkvenjakov R, Dickson M, Drmanac R, Drmanac S;



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 07:13:00 ; Search time 3518.14 Seconds

(without alignments)  
11357.738 Million cell updates/sec

Title: US-09-245-198a-3

Perfect score: 1373

Sequence: 1 atgcattgttagcttga.....gacaaatgtgataatg 1373

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBank: \*  
1: gb\_ba: \*  
2: gb\_htg: \*  
3: gb\_in: \*  
4: gb\_om: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sts: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*  
15: em\_ba: \*  
16: em\_fun: \*  
17: em\_hum: \*  
18: em\_in: \*  
19: em\_mu: \*  
20: em\_om: \*  
21: em\_or: \*  
22: em\_ov: \*  
23: em\_pat: \*  
24: em\_ph: \*  
25: em\_pl: \*  
26: em\_ro: \*  
27: em\_sts: \*  
28: em\_un: \*  
29: em\_vl: \*  
30: em\_htg\_hum: \*  
31: em\_htg\_inv: \*  
32: em\_htg\_other: \*  
33: em\_htg\_mus: \*  
34: em\_htg\_pln: \*  
35: em\_htg\_rtd: \*  
36: em\_htg\_mam: \*  
37: em\_htg\_vrl: \*  
38: em\_sy: \*  
39: em\_htgo\_hum: \*  
40: em\_htgo\_mus: \*  
41: em\_htgo\_other: \*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1330.2	96.2	1353	6	AX201324	AX201324 Sequence
2	1320.2	96.2	1368	9	AF055872	AF055872 Homo sapi
3	1285	93.6	1306	9	AF030099	AF030099 Homo sapi
4	1226.4	89.3	1236	6	AR140407	AR140407 Sequence
5	1096.8	79.9	1651	9	BC019047	BC019047 Homo sapi
6	768.2	56.0	177703	2	AC016876	AC016876 Homo sapi
7	754	54.9	215795	2	AC127470	AC127470 pan trogl
8	629.2	45.8	898	6	AX180714	AX180714 Sequence
9	614.6	44.8	1168	10	AF030100	AF030100 Mus muscu
10	364.2	26.5	15353	2	AC126921	AC126921 Bos tauru
11	309.6	22.5	203083	2	AC069459	AC069459 Mus muscu
12	309.6	22.5	234182	10	AL603707	AL603707 Mouse DNA
13	254.2	18.5	138792	2	AC119115	AC119115 Rattus no
14	246.4	17.9	177555	2	AC130192	AC130192 Sus scrofa
15	242.4	17.7	161428	2	AC126925	AC126925 Canis fam
16	193.4	14.1	195	6	AX379024	AX379024 Sequence
17	158.6	11.6	184026	2	AC098923	AC098923 Rattus no
18	111	8.1	203281	2	AC126237	AC126237 Felis cat
19	100.6	7.3	139405	2	AC126239	AC126239 Felis cat
20	87.2	6.4	7218	6	AF429315	AF429315 Sequence 14
21	77.6	5.7	125020	9	AF429315	AF429315 Homo sapi
22	74.8	5.4	303091	2	AC084799	AC084799 Mus muscu
23	71	5.2	121451	2	AC111931	AC111931 Rattus no
24	70.8	5.2	73948	2	AC022556	AC022556 Homo sapi
25	70.6	5.1	144979	2	AC016280	AC016280 Homo sapi
26	70.4	5.1	220469	2	AC074307	AC074307 Mus muscu
27	70.2	5.1	298166	2	AC087563	AC087563 Homo sapi
28	69.2	5.0	.936	8	CNS01AYK	AL113556 Botrytis
29	69.2	5.0	82897	2	AC022595	AC022595 Homo sapi
30	68.4	5.0	202083	2	AC023833	AC023833 Mus muscu
31	68.2	5.0	110737	2	AC011105	AC011105 Homo sapi
32	68.2	5.0	265537	2	AC087228	AC087228 Mus muscu
33	68	5.0	181344	2	AC125880	AC125880 Rattus no
34	68	5.0	300695	2	AC079431	AC079431 Mus muscu
35	67.6	4.9	626	8	CNS018RP	AL110716 Botrytis
36	67.2	4.9	66624	2	AC101520	AC101520 Mus muscu
37	67.2	4.9	147124	2	AC022768	AC022768 Homo sapi
38	67.2	4.9	177883	2	AC046159	AC046159 Homo sapi
39	67.2	4.9	181988	2	AC090552	AC090552 Homo sapi
40	67.2	4.9	252689	2	AC079433	AC079433 Homo sapi
41	67	4.9	74138	2	AC021272	AC021272 Homo sapi
42	66.8	4.9	62649	2	AC022552	AC022552 Homo sapi
43	66.8	4.9	75628	2	AC021793	AC021793 Homo sapi
44	66.8	4.9	100511	2	AC010774	AC010774 Homo sapi
45	66.8	4.9	205691	2	AC087227	AC087227 Mus muscu

#### ALIGNMENTS

RESULT 1	AX201324	1353 bp	DNA	Linear	PAT 30-AUG-2001
LOCUS	AX201324				
DEFINITION	Sequence 3 from Patent WO0153486.				
ACCESSION	AX201324				
VERSION	AX201324.1	GI:15391154			
KEYWORDS					
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1353)				
AUTHORS	Ashkenazi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L., Hillan,K.J., Marsters,S.A., Pan,J., Pittl,R.M., Roy,M.A., Smith,V., Stone,D.M., Watanabe,C.K. and Wood,W.I.				

OY	49	CGATTCCTCCTGGGTTCCCGGGATGGGGGGGGGTGTAGACGACGACACCCCCCCCATTG	1
Db	1	CGATCCCTCGGGTCCCGGGATGGGGGGGGGTGTAGACGACGACACCCCCCCCATTG	60
OY	109	GCCGCCCTTCGGAACCAAGCCGGAGGGGGCGCCGGGGGGAGCCGGACCCGCCTGTG	168
Db	61	GCCGCCCTTCGGAAGCAGAGGGGGGAGGGGGGGCCGGGGGGAGCCGGGACCCGCTGTG	121
OY	169	GTCCCGCTCGSCTGGGCGCTGGGGCTGGCGGTGGCTGCTCGGCTCTCTGTGGCGTG	222
Db	121	GTCCCGCTCGSCTGGGCGCTGGGGCTGGCGGTGGCTGCTCGGCTCTCTGTGGCGTG	180
OY	229	GTCACTTTGGGAGCCGGGCACTGCTGTCCGCCAGAGACCCTGCCCAGAGAGACTGGTG	288
Db	181	GTCACTTTGGGAGCCGGGCACTGCTGTCCGCCAGAGACCCTGCCCAGAGAGACTGGTG	241
OY	289	GCAGAGAGAGACAGAGACCCTGTCCGAATAATCCCGAAGAAAGAAAGCAGATCT	344
Db	241	GCAGAGAGAGACAGAGACCCTGTCCGAATAATCCCGAAGAAAGAAAGCAGATCT	300
OY	349	GCGCCTTCTCTGAACCGACTAGTTGCGCCTGSCAGAAATGCACCTTAAAGCCCGAAAAA	408
Db	301	GCGCCTTCTCTGAACCGACTAGTTGCGCCTGSCAGAAATGCACCTTAAAGCCCGAAAAA	366
OY	409	CGGGCTGGAAGGGGATGGGAGCCCATTAATAATTATCTCACACTGTGACAGAGACGA	466
Db	361	CGGGCTGGAAGGGGATGGGAGCCCATTAATAATTATCTCACACTGTGACAGAGACGA	420
OY	469	GCGCAGCAGGTTGTGGACGGGACAGTAGTAGTGGCTGGGGAGAGCAGAATTAACAGCTCC	528
Db	421	GCGCAGCAGGTTGTGGACGGGACAGTAGTAGTGGCTGGGGAGAGCAGAATTAACAGCTCC	480
OY	529	AGCCCTCTGGCCTTCAACCCGCAATGGGGAGATTATAGTCAACCGGGCTGGGCTCTAC	588
Db	481	AGCCCTCTGGCCTTCAACCCGCAATGGGGAGATTATAGTCAACCCGGGCTGGGCTCTAC	540
OY	589	TACCTGTACTCTCAGGTGCATTTGTATGAGGGAGAAAGCTGTCTACCTGMAAGCTGAACTTG	648
Db	541	TACCTGTACTCTCAGGTGCATTTGTATGAGGGAGAAAGCTGTCTACCTGMAAGCTGAACTTG	600
OY	649	CTGGTGGATGGTGTGTGGGCCCTGCGCTGCCGTGGAGAAATTCAGCCACTGCGGCGAGT	708
Db	601	CTGGTGGATGGTGTGTGGGCCCTGCGCTGCCGTGGAGAAATTCAGCCACTGCGGCGAGT	660
OY	709	TCCCTGGGAGCCCAAGCTCCGCGCTGTGCAGAGTGTGGGCTGTGGGCCCTGCGGCAAGG	768
Db	661	TCCCTGGGAGCCCAAGCTCCGCGCTGTGCAGAGTGTGGGCTGTGGGCCCTGCGGCAAGG	720
OY	769	TCTTCCTCGGAGTCGACACCTCCCTCGGAGCCCATTCGAAGCTGCCCCCTTCTCTACC	828
Db	721	TCTTCCTCGGAGTCGACACCTCCCTCGGAGCCCATTCGAAGCTGCCCCCTTCTCTACC	780
OY	829	TACTTGGAGACTCTCAAGATTCACTAGAGGGCCCTGCTCTCCCAACAGTGTCCAAGCT	888
Db	781	TACTTGGAGACTCTCAAGATTCACTAGAGGGCCCTGCTCTCCCAACAGTGTCCAAGCT	840
OY	889	GCGGCTTCCCTTCAGACAGCTCTGTGGGACACCGGCTCCCTCTGCCCCACACTAGGCGGT	948
Db	841	GCGGCTTCCCTTCAGACAGCTCTGTGGGACACCGGCTCCCTCTGCCCCACACTAGGCGGT	900

OY	949	CTTTGGTCAGAGCTGCCCCCTCCCTCTAGAGGCTGCGCTGGACCTTTACAGATTTTCCA	1008
Db	901	CTTTGGTCCAGACGTGCCCCCTCCCTCTAGAGGCTGCGCTGGACCTTTTCCA	960
OY	1009	TTCCCACTAATAATACAGTATTTCCCACTCTTATCTTACAACTCCCCACGCCCACTCTCCA	1068
Db	961	TTCCCACTAATAATAAGTATTTCCCACTCTTATCTTACAACTCCCCACGCCCACTCTCCA	1020
OY	1069	CCTCACTAGCTCCCAATAATCCCTGACCCCTTGAGGGCCCCAGTAGATCTGACATCCCCCTGG	1128
Db	1021	CCTCACTAGCTCCCAATAATCCCTGACCCCTTGAGGGCCCCAGTAGATCTGACATCCCCCTGG	1080
OY	1129	GCCACAGACCCCCAGGCGCATTTGTGTCACTACTCTGTGGCGAAGATGGGTCCAGAAG	1188
Db	1081	GCCACAGACCCCCAGGCGCATTTGTGTCACTACTCTGTGGCGAAGATGGGTCCAGAAG	1140
OY	1189	ACCCCACTTCAGGCACTAAGAGGGGCTGAGCTTGGCGCGAGAGGCCAAGAAGACTGGGC	1248
Db	1141	ACCCCACTTCAGGCACTAAGAGGGGCTGAGCTTGGCGCGAGAGGCCAAGAAGACTGGGC	1200
OY	1249	CTAGGCCAGAGGATGCCAAATGTGAGGGGCGGAAACAAGAAAGCTCTCCCTTGAGAA	1308
Db	1201	CTAGGCCAGAGGATGCCAAATGTGAGGGGCGGAAACAAGAAAGCTCTCCCTTGAGAA	1260
OY	1309	TTCCCTGTGGAATTTTAAACAGATATTTATTTATTTATTTGTGACAAATGTTTGATA	1368
Db	1261	TTCCCTGTGGAATTTTAAACAGATATTTATTTATTTATTTGTGACAAATGTTTGATA	1320
OY	1369	AATGG 1373	
Db	1321	AATGG 1325	

RESULT	2
LOCUS	AF055872
DEFINITION	Homo sapiens Apo3/DR3 ligand (APO3L) mRNA, complete cds.
VERSION	AF055872
KEYWORDS	.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1368) Marshers,S.A., Sheridan,J.P., Pittl,R.M., Brush,J., Goddard,A. and Ashkenazi,A. Identification of a ligand for the death-domain-containing receptor Apo3 Curr. Biol. 8 (9), 525-528 (1998) 98228355 9560343 2 (bases 1 to 1368) Marshers,S.A., Sheridan,J.P., Pittl,R.M., Brush,J., Goddard,A. and Ashkenazi,A. Direct Submission Submitted (25-MAR-1998) Molecular Oncology, Genentech, 1 DNA Way, South San Francisco, CA 94080, USA location/Qualifiers 1..1368 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="17" /map="17p13" 1..1368 /gene="APO3L" 58..807 /gene="APO3L" /function="binds to the death-domain-containing receptor Apo3/DR3" /note="TNF homolog" /codon_start=1
TITLE	JOURNAL
AUTHORS	REFERENCE
PUBMED	JOURNAL
FEATURES	source
gene	gene
CDS	CDS

/product="Ap03/DR3 11gand"  
 /protein\_id="AAC39724.1"  
 /db\_xref="GI:3108231"  
 /translation="MAARSOBRNRGRREPGNALLVPLALGLIALACGLILAVSL  
 GSRASLSHQERAOELVABEDDDPELNPOTRESDPAPFLKRLVPRPSAFKRTK  
 ARRAIAAHYEVHPRPDQAGVDGIVSGWEAEAINSSPLRYNRQIGEFVITAGL  
 YLLCQVHFDEBKAAVYLKLDLVDVLAIRCRIEESATSAASLSIGQLRQVSGLLAT  
 RPSLSLRRTTPMHLKAPFLTYFLGVH"  
 BASE COUNT 272 a 443 c 389 g 264 t  
 ORIGIN

Query Match 96.2%; Score 1320.2; DB 9; Length 1368;  
 Best Local Similarity 99.8%; Pred. No. 7.9e-231;  
 Matches 1322; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 49 CGATCCCTCGGGTCCGGGGTGGGGGGGGGGTGGAGGACAGAGCCCGCCCGCCGATG 108  
 DB 1 CGATCCCTCGGGTCCGGGGTGGGGGGGGGGTGGAGGACAGAGCCCGCCCGCCGATG 60  
 QY 109 GCGGCCCTCGGAGCCAGAGGCGGAGGGGGGCGCCGGGGGAGCCGGGCAACCGCTGCTG 168  
 DB 61 GCGGCCCTCGGAGCCAGAGGCGGAGGGGGGCGCCGGGGGAGCCGGGCAACCGCTGCTG 120  
 QY 169 GTCCCGCTCGCGCTGGGCTGGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 228  
 DB 121 GTCCCGCTCGCGCTGGGCTGGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 180  
 QY 229 GTCCAGTTGGGGAGCCGGGCGATCGCTGCGCCGACAGAGCCCGCCCGCCGAGAGCTGGT 288  
 DB 181 GTCCAGTTGGGGAGCCGGGCGATCGCTGCGCCGACAGAGCCCGCCCGCCGAGAGCTGGT 240  
 QY 289 GCAGAGAGAGAGACAGAGCCCGTGGAACTGAATCCCGACAGAGAGAGAGAGAGAGTCT 348  
 DB 241 GCAGAGAGAGAGACAGAGCCCGTGGAACTGAATCCCGACAGAGAGAGAGAGAGTCT 300  
 QY 349 GCGCCTTCTCTGACCGAGTCTGGGCTCGCAGAGAGTGCACCTTAAGGCGGAAAGCA 408  
 DB 301 GCGCCTTCTCTGACCGAGTCTGGGCTCGCAGAGAGTGCACCTTAAGGCGGAAAGCA 360  
 QY 409 CGGGCTCGAAGAGCGATGCGAGCCCATTTATGAATTCATCCAGAGAGAGAGAGAGAG 468  
 DB 361 CGGGCTCGAAGAGCGATGCGAGCCCATTTATGAATTCATCCAGAGAGAGAGAGAGAG 420  
 QY 469 GCGCAGAGAGAGTGTGAGACGAGTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGTCT 528  
 DB 421 GCGCAGAGAGAGTGTGAGACGAGTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGTCT 480  
 QY 529 AGCCCTCTGCGCTACACCGCCAGATCGGGAGATTATAGTCACCGGGCTGGGCTCTAC 588  
 DB 481 AGCCCTCTGCGCTACACCGCCAGATCGGGAGATTATAGTCACCGGGCTGGGCTCTAC 540  
 QY 589 TACCTGACTGTCAGGTGACCTTGTATGAGGGGAGAGGCTGTACTGAAAGTGGAGCTTG 648  
 DB 541 TACCTGACTGTCAGGTGACCTTGTATGAGGGGAGAGGCTGTACTGAAAGTGGAGCTTG 600  
 QY 649 CTGCTGATGATGTGCTGCGCTGCGCTGCGTGGAGAGATTCTACAGCACTGCGGCGAGT 708  
 DB 601 CTGCTGATGATGTGCTGCGCTGCGCTGCGTGGAGAGATTCTACAGCACTGCGGCGAGT 660  
 QY 709 TCCCTGAGGAGCCAGTCCGCTCTGCGAGGTGTCTGGGCTGTGGCCCTGCGGCGAGG 768  
 DB 661 TCCCTGAGGAGCCAGTCCGCTCTGCGAGGTGTCTGGGCTGTGGCCCTGCGGCGAGG 720  
 QY 769 TCCCTGAGGAGCCAGTCCGCTCTGCGAGGTGTCTGGGCTGTGGCCCTGCGGCGAGG 828  
 DB 721 TCCCTGAGGAGCCAGTCCGCTCTGCGAGGTGTCTGGGCTGTGGCCCTGCGGCGAGG 780  
 QY 829 TACTTGAGACTCTTCCAGGTTCACTGAGGGGCGCTGCTCCCGACAGTGTCCAGGCT 888  
 DB 781 TACTTGAGACTCTTCCAGGTTCACTGAGGGGCGCTGCTCCCGACAGTGTCCAGGCT 840  
 QY 889 GCGGCTCCCTCGAGAGCTCTGTGGGACCGGCTCCCTCTGCGCCAGCCCTGAGCGCT 948

DB 841 GCGGCTCCCTCGAGAGCTCTGTGGGCAACCCGGTCCCTCTGCGCCCGCCAGCCGCT 900  
 QY 949 CTGTGTCAGAGCTGCGCCCTCCCTCTAGAGGCTGCGGCTGCTGTCAGGTCTTCA 1008  
 DB 901 CTGTGTCAGAGCTGCGCCCTCCCTCTAGAGGCTGCGGCTGCTGTCAGGTCTTCA 960  
 QY 1009 TCCCATATAATACAGTATTCACCTTATCTTATCTTACACTCCCGCCAGCTCTCA 1068  
 DB 961 TCCCATATAATACAGTATTCACCTTATCTTATCTTACACTCCCGCCAGCTCTCA 1020  
 QY 1069 CCTCATAGTCTCCCAATCCCTGACCTTTAGGCGCCCGCCAGTATCTGACTCCCTG 1128  
 DB 1021 CCTCATAGTCTCCCAATCCCTGACCTTTAGGCGCCCGCCAGTATCTGACTCCCTG 1080  
 QY 1129 GGCACAGACCCCGAGGCAATGTGTCACGTCTGTCGTGGGCAAGAGAGTGGTCAAG 1188  
 DB 1081 GGCACAGACCCCGAGGCAATGTGTCACGTCTGTCGTGGGCAAGAGAGTGGTCAAG 1140  
 QY 1189 ACCCACTTCAGGCACTAAGAGGGGCTGACCTGCGGCGAGAGAGCAAGAGACTGGG 1248  
 DB 1141 ACCCACTTCAGGCACTAAGAGGGGCTGACCTGCGGCGAGAGAGCAAGAGACTGGG 1200  
 QY 1249 CTAGGCGAGAGTTCCTCAATGTGAGGGGCGAGAAACAGACAGCTCTCTCTAGAA 1308  
 DB 1201 CTAGGCGAGAGTTCCTCAATGTGAGGGGCGAGAAACAGACAGCTCTCTCTAGAA 1260  
 QY 1309 TTCCCTGTGATTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTT 1368  
 DB 1261 TTCCCTGTGATTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTT 1320  
 QY 1369 AATGG 1373  
 DB 1321 AATGG 1325

RESULT 3  
 AF030099 1306 bp mRNA 1linear PRI 20-DEC-1997  
 LOCUS  
 DEFINITION  
 ACCESSION AF030099  
 VERSION AF030099.1 GI:2707218  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE  
 AUTHORS  
 TITLE  
 Chicheportliche, Y., Bourdon, P.R., Xu, H., Hsu, Y.M., Scott, H.,  
 Hession, C., Garcia, I., and Browning, J. L.  
 TWEAK, a new secreted ligand in the tumor necrosis factor family  
 that weakly induces apoptosis  
 J. Biol. Chem. 272 (51), 32401-32410 (1997)  
 JOURNAL  
 MEDLINE  
 PUBMED 9405449  
 2 (bases 1 to 1306)  
 REFERENCE  
 AUTHORS  
 TITLE  
 Bourdon, P., Hession, C., Tizard, R. and Browning, J.  
 Direct Submission  
 Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,  
 Cambridge, MA 02142, USA  
 JOURNAL  
 FEATURES  
 source

1.1306  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /chromosome="17"  
 /map="17p13"  
 /tissue\_type="tonsil"  
 18.767  
 /note="ligand in the TNF family; secreted protein; start  
 codon not verified experimentally"  
 /codon\_start=1  
 /product="TWEAK"  
 /protein\_id="AAC51923.1"  
 /db\_xref="GI:2707219"

## CDS

889 GCGGCTCCCTCGAGAGCTCTGTGGGACCGGCTCCCTCTGCGCCAGCCCTGAGCGCT 948

/translation-"MAARRSORRRRNGEPTALIVPLALGLALACLLILAVSL  
 GSRASISADPEADELVAEDQDPSLNPOTSESDAPLNLVRNRSAPKGRTR  
 ARRAIAHYEHPGQDAGADVTGDSGMEERIDINSSPLRNRIGETITRRL  
 YLYVCQHYDEGKAVYKLDLVDGVALCLIEFSATIASLSLQPLRLCOVSLIAL  
 RPSLSLIRPLNHLKAKAPFLYFGLFYH"  
 BASE COUNT 247 a 434 c 368 g 257 t  
 ORIGIN

Query Match 93.6%; Score 1285; DB 9; Length 1306;  
 Best Local Similarity 100.0%; Pred. No. 2,1e-224;  
 Matches 1285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 89 CACAGCCCCCCCCCCCCCTGGCCCGCTCGAGCCAGAGCCGAGAGGGGGCCGGGGGG 148  
 DB 1 CACAGCCCCCCCCCCCCCTGGCCCGCTCGAGCCAGAGCCGAGAGGGGGCCGGGGGG 60  
 QY 149 AGCGGCGACCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 208  
 DB 61 AGCGGCGACCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
 QY 209 TCGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 268  
 DB 121 TCGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 QY 269 CTGGCCAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 328  
 DB 181 CTGGCCAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 QY 329 CAGAAGAAAGCCAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 388  
 DB 241 CAGAAGAAAGCCAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
 QY 389 CACCTAAAGCCGAGAAACAGAGGCTGGAAGAGGATCGAGCCCATTTATGAGTTTCATC 448  
 DB 301 CACCTAAAGCCGAGAAACAGAGGCTGGAAGAGGATCGAGCCCATTTATGAGTTTCATC 360  
 QY 449 CACGACCTGACAG 508  
 DB 361 CACGACCTGACAG 420  
 QY 509 AAGCCAGAAATCAACAGCTCCAGCCCTCTGCGCTACAAACCCGAGATCGGGAGTTTATG 568  
 DB 421 AAGCCAGAAATCAACAGCTCCAGCCCTCTGCGCTACAAACCCGAGATCGGGAGTTTATG 480  
 QY 569 TCACCCGGGCTGGGCTCTACTACTGTAAGTGTGCACTTTGATGAGGGAGAGGCTG 628  
 DB 481 TCACCCGGGCTGGGCTCTACTACTGTAAGTGTGCACTTTGATGAGGGAGAGGCTG 540  
 QY 629 TCTACTGAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 688  
 DB 541 TCTACTGAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 QY 689 TCTGAGCACTGCGGCGAGTTCCCTCGGGGCCAGCTCCGCTCTGCGAGGTGTCTGGGC 748  
 DB 601 TCTGAGCACTGCGGCGAGTTCCCTCGGGGCCAGCTCCGCTCTGCGAGGTGTCTGGGC 660  
 QY 749 TGTGTGGCCCTGCGGCGAGGTTCTCCCTGGGAGATCCGACCCCTCCCTGGGCCATCTCA 808  
 DB 661 TGTGTGGCCCTGCGGCGAGGTTCTCCCTGGGAGATCCGACCCCTCCCTGGGCCATCTCA 720  
 QY 809 AGGCTGCCCCCTTCTGACCTACTTGGAGCTCTTCCAGATTCAATGAGGGGCCCTGTGCT 868  
 DB 721 AGGCTGCCCCCTTCTGACCTACTTGGAGCTCTTCCAGATTCAATGAGGGGCCCTGTGCT 780  
 QY 869 CCCCACAGTGTCCAGAGCTGCGGCTCCCTCGACAGCTCTTGGGACACCCGGTCCCT 928  
 DB 781 CCCCACAGTGTCCAGAGCTGCGGCTCCCTCGACAGCTCTTGGGACACCCGGTCCCT 840  
 QY 929 CTGCGCCACCTCTGCGGCTCTTGTCTCAGAGACTGCCCCCTCCCTAGAGAGCTGCTG 988  
 DB 841 CTGCGCCACCTCTGCGGCTCTTGTCTCAGAGACTGCCCCCTCCCTAGAGAGCTGCTG 900  
 QY 989 GCCTGTTCAGCTGTTTTCATCCACATTAATACAGTATTCACCTTATCTTACAAT 1048

DB 901 GCTGTTCACGCTGTTTTCATCCACATTAATACAGTATTCACCTTATCTTACAAT 960  
 QY 1049 CCCCACAGGCGACCTCTGACCTACTAGTCCCAATCCCGTACCCTTTAGGCCCCCA 1108  
 DB 961 CCCCACAGGCGACCTCTGACCTACTAGTCCCAATCCCGTACCCTTTAGGCCCCCA 1020  
 QY 1109 GTGATCTGACCTCCCTGCGCACAGACCCCAAGGCGATGTCTTACTGACTGCTG 1168  
 DB 1021 GTGATCTGACCTCCCTGCGCACAGACCCCAAGGCGATGTGTCTTACTGCTG 1080  
 QY 1169 GCGAAGATGGGTCCAGAAACCCCACTTACGCACTAAGAGGGGCTGACCTGGCGCA 1228  
 DB 1081 GCGAAGATGGGTCCAGAAACCCCACTTACGCACTAAGAGGGGCTGACCTGGCGCA 1140  
 QY 1229 GGAAGCCAAAGAGAGCTGAGGCGCAGAGATTCCCAATGTGAGGGGCGAGAAACAG 1288  
 DB 1141 GGAAGCCAAAGAGAGCTGAGGCGCAGAGATTCCCAATGTGAGGGGCGAGAAACAG 1200  
 QY 1289 ACAAGCTCTCTCTGAGAAATTCCTGTGATTTTAAACAGATATTTATTTATTT 1348  
 DB 1201 ACAAGCTCTCTCTGAGAAATTCCTGTGATTTTAAACAGATATTTATTTATTT 1260  
 QY 1349 ATTGTGACAAATGTTGATTAATG 1373  
 DB 1261 ATTGTGACAAATGTTGATTAATG 1285

RESULT 4  
 ARI40407 1236 bp DNA linear PAT 16-JUN-2001  
 LOCUS Sequence 1 from patent US 6207642.  
 DEFINITION ARI40407  
 ACCESSION ARI40407  
 VERSION ARI40407.1 GI:14482903  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1236)  
 AUTHORS Wiley,S.R.  
 TITLE Member of the TNF family useful for treatment and diagnosis of disease

JOURNAL Patent: US 6207642-A 1 27-MAR-2001;  
 FEATURES Location/Qualifiers  
 source 1..1236  
 BASE COUNT 225 a 416 c 358 g 237 t

Query Match 89.3%; Score 1226.4; DB 6; Length 1236;  
 Best Local Similarity 99.5%; Pred. No. 1e-213;  
 Matches 1230; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 106 ATGGCCGCCCTCGGACCCAGAGCGAGAGGGGGCGGGGGAGCGGGGACCGCCCTG 165  
 DB 1 ATGGCCGCCCTCGGACCCAGAGCGAGAGGGGGCGGGGGAGCGGGGACCGCCCTG 60  
 QY 166 CTGCTCCCGCTGCGGCTGAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 225  
 DB 61 CTGCTCCCGCTGCGGCTGAGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 120  
 QY 226 GTGTGAGTTTGGGAGCCCGGCAATCGCTGTCCGCCAGAGAGCTGGCCAGAGAGACTG 285  
 DB 121 GTGTGAGTTTGGGAGCCCGGCAATCGCTGTCCGCCAGAGAGCTGGCCAGAGAGACTG 180  
 QY 286 GTGGCAAGAGAGACAGAGACCGCTCGGAATGAATCCCAAGAGAGAAAGCAGAGAT 345  
 DB 181 GTGGCAAGAGAGACAGAGACCGCTCGGAATGAATCCCAAGAGAGAAAGCAGAGAT 240  
 QY 346 CTGCGGCTTCTCGAAGCCGACTAGTTGGGCTGCGAAGTGAAGTGAAGGCGGGAAA 405  
 DB 241 CTGCGGCTTCTCGAAGCCGACTAGTTGGGCTGCGAAGTGAAGTGAAGGCGGGAAA 300

OY	406	ACACGGGCTCCAGAAAGGCGATCGCAGGCCCATATATGAGATTCATCCACGACCTGAGCAAGAC	465
Db	301	ACACGGGCTCCAGAAAGGCGATCGCAGGCCCATATATGAGATTCATCCACGACCTGAGCAAGAC	360
OY	466	GGAGGCGCAGGCGAGGTGTGGACGGGACACTGATGATGGCTGGAGGAAAGCAGATCAACAGC	525
Db	361	GGAGGCGCAGGCGAGGTGTGGACGGGACACTGATGATGGCTGGAGGAAAGCAGATCAACAGC	420
OY	526	TCACAGCCCTTGCGCTACACCGCCAGATCGGGGAGTTTATATGATCACCCGGGCTGGGCTC	585
Db	421	TCACAGCCCTTGCGCTACACCGCCAGATCGGGGAGTTTATATGATCACCCGGGCTGGGCTC	480
OY	586	TACATCCGTACTGTGACAGTGCATTTATATAGGGGAAAGGCTGCTACTGTAAGCTGGAC	645
Db	481	TACATCCGTACTGTGACAGTGCATTTATATAGGGGAAAGGCTGCTACTGTAAGCTGGAC	540
OY	646	TTGCGTGTGATATGATGCTGGCCCTGGCGCTCCGAGAGAAATTCAGACCACCTGGGCC	705
Db	541	TTGCGTGTGATATGATGCTGGCCCTGGCGCTCCGAGAGAAATTCAGACCACCTGGGCC	600
OY	706	AGTTCCCTCGGGGCCCCAGACTCCGCTCTGCCAGGTGTCTGGGCTGTTGAGCCCTGGGGCCA	765
Db	601	AGTTCCCTCGGGGCCCCAGACTCCGCTCTGCCAGGTGTCTGGGCTGTTGAGCCCTGGGGCCA	660
OY	766	GGGTCTCTCCCTCGGGATCCGACACCTCCCTTGGGCCATCTCAGAGGCTGCCCTTCTCTC	825
Db	661	GGGTCTCTCCCTCGGGATCCGACACCTCCCTTGGGCCATCTCAGAGGCTGCCCTTCTCTC	720
OY	826	ACCTACTTCGGGACTTTCAGAGTCTAGAGGGGCCCTGGTCCGCCACAGTCGTCAG	885
Db	721	ACCTACTTCGGGACTTTCAGAGTCTAGAGGGGCCCTGGTCTCCCGCAGTCGTCAG	780
OY	886	GCTGCGGAGCTCCCTCGACAGCTCTGTGGGACACCGGTCCTCTGCCCCACCCCTCAGCC	945
Db	781	GCTGCGGAGCTCCCTCGACAGCTCTGTGGGACACCGGTCCTCTGCCCCACCCCTCAGCC	840
OY	946	GCTTTTGTCTCAGACGTCGCCCTCCCTCTAGAGGCTGCTGGGGCTGTTACGATGTTTT	1005
Db	841	GCTTTTGTCTCAGACGTCGCCCTCCCTCTAGAGGCTGCTGGGGCTGTTACGATGTTTT	900
OY	1006	CCATCCCAATAAATACAGTATTCACACTTATTTCAACTCCGCCACGCCCACTCT	1065
Db	901	CCATCCCAATAAATACAGTATTCACACTTATTTCAACTCCGCCACGCCCACTCT	960
OY	1066	CCACTCTACACTGCTCCCAATCTCTGACCCCTTTGAGGCCCCCAAGTATCTGACTCCCC	1125
Db	961	CCACTCTACACTGCTCCCAATCTCTGACCCCTTTGAGGCCCCCAAGTATCTGACTCCCC	1020
OY	1126	CTGGGCAAGAACCCCAAGGCGATTGTGTACTGTACTGTACTGTGTGGGCAAGGATGGGTCAG	1185
Db	1021	CTGGGCAAGAACCCCAAGGCGATTGTGTACTGTACTGTACTGTGTGGGCAAGGATGGGTCAG	1080
OY	1186	AAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCGAGAGACCAAGAGACTG	1245
Db	1081	AAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGCGAGAGACCAAGAGACTG	1140
OY	1246	GGCTTAGGCGCAGAGTCCCAATGTGAGGGGCGAAGAACAGACAACTCTCTCTCTGA	1305
Db	1141	GGCTTAGGCGCAGAGTCCCAATGTGAGGGGCGAAGAACAGACAACTCTCTCTCTGA	1200
OY	1306	GAATTCCTCTGTGATTTTAAACAGATATTTT 1341	
Db	1201	GAATTCCTCTGTGATTTTAAACAGATATTTT 1236	

RESULT 5	1651 bp	mRNA	linear	PRI 11-DEC-2001
BC019047				
LOCUS				
DEFINITION	Homo sapiens, Similar to tumor necrosis factor (ligand)			
	superfamily, member 12, clone MGC:20669			
	complete cds.			
ACCESSION	BC019047			
VERSION	BC019047.1	GI:17512138		

KEYWORDS	MGC.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1 (bases 1 to 1651)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL.	Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk

Email: [cgabbs@mail.nih.gov](mailto:cgabbs@mail.nih.gov)  
Tissue Procurement: Louis Staudt  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNI)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc.mcgc@nih.gov](mailto:nisc.mcgc@nih.gov)  
Shaychenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,  
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
Lim, M., Maduro, Q.L., Mastaglio, C., Mastrian, S.D., McCloskey, J.C.,  
McConnell, J., Pearson, R., Snyder, B., Statnikov, S., Thomas, P.J.,  
Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,  
Zhang, L., H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 30 Row: P Column: 5  
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507596.  
 Location/Qualifiers  
 1..1651

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_MGC="20669 IMAGE:4766071"
/tissue_type="Primary B-Cells from Tonsils"
/clone_id="NH_MGC_48"
/lab_host="DH10B-R"
/clone_vector="pOTB7"
106. .510

```

```

/codon_start=1
/product="Similar to tumor necrosis factor (ligand)
superfamily, member 12"
/protein_id="AAH19047.1"
/db_xref="GI:17512139"
/translation="MAARSRQRRGRGEPETALLPLALGIALACLGILLAVSL
GSRSLAEOEPAOEELVAEDODPSEINPQTEESQDPAPFLNRLVRRBSRAPKGRKTR
ARRRLAAHYEHVHRRPGDGAADGGYTTCLRP"

```

Query Match	79.9%	Score 1096.8	DB 9	Length 1651
Best Local Similarity	84.8%	Pred. No. 4.1e-190		
Matches 1347	Conservative	0	Mismatches 2	Indels 239
			Gaps	1
QY	25	TCGCGCGCGGCTCCCTCCCGATCCCTCGGGTCCCGGAGGGGGCGGTGAG	84	
DB	25	TCGCGCGCGGCTCCCTCCCGATCCCTCGGGTCCCGGAGGGGGCGGTGAG	84	
QY	85	CAGGCACAGCCCGCGCCCATGGCCGCGTCGAGACAGAGCGAGGGGCGCGG	144	
DB	85	CAGGCACAGCCCGCGCCCATGGCCGCGTCGAGACAGAGCGAGGGGCGCGG	144	
QY	145	GGGGAGCCGGGACCGGCTGCTGGTCCCGCTCGGGCTGAGGCTTGGGCTGGCC	204	
DB	145	GGGGAGCCGGGACCGGCTGCTGGTCCCGCTCGGGCTGAGGCTTGGGCTGGCC	204	



Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
Testfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
Zembock, L., Zimmer, A. and Zody, M.

**TITLE** Direct Submission  
**JOURNAL** Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genomics Research, 320 Charles Street, Cambridge, MA 02141, USA  
**COMMENT** On Jun 2, 2002 this sequence version replaced g1:15421989.

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WMRB  
 Web site: <http://www.seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 -----  
 Project Information  
 Center project name: L2849  
 Center clone name: 186\_B\_7

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 10 'contigs'. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```

* 1 46645: contig of 48645 bp in length
* 48656 48745: gap of 100 bp
* 48746 58277: contig of 9532 bp in length
* 58278 58377: gap of 100 bp
* 58378 65804: contig of 7477 bp in length
* 65805 65904: gap of 100 bp
* 65905 79793: contig of 13689 bp in length
* 79794 79893: gap of 100 bp
* 79894 99493: contig of 19600 bp in length
* 99494 99593: gap of 100 bp
* 99594 111049: contig of 11456 bp in length
* 11050 11149: gap of 100 bp
* 11150 123020: contig of 13871 bp in length
* 125021 125150: gap of 100 bp
* 125151 145109: contig of 19889 bp in length
* 145110 145209: gap of 100 bp
* 145210 166458: contig of 24249 bp in length
* 166459 169558: gap of 100 bp
* 169559 177703: contig of 8145 bp in length

```

FEATURES	Location/Qualifiers
source	1. .177703

[illegible]

Query Match	56.0%	Score 768.2	DB 2	Length 177703
Best Local Similarly	99.6%	Pred. No. 2.4e-130		
Matches 770; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	601	CAGGTGCACATTGATGAGGGGAGAGCGCTGTCTACCTGAAAGCTGGACTGTGCTGGATGGT	660
Db	37314	CAGGTGCACATTGATGAGGGGAGAGCGCTGTCTACCTGAAAGCTGGACTGTGCTGGATGGT	37255
QY	661	GTGCTGGCCCTGCGCTGTGGCTGGAGAAATTCATAGGCACATGGGGCAGTTCCTCGGGGCC	720
Db	37254	GTGCTGGCCCTGCGCTGTGGCTGGAGAAATTCATAGGCACATGGGGCAGTTCCTCGGGGCC	37195
QY	721	CAGCTCCGCTTGTGCAGGTGTCTGGGCTGTGGCCCTGCGGCACAGAGTCTCTCCCTGCGG	780
Db	37194	CAGCTCCGCTTGTGCAGAGTGTCTGGTGTGTGGGCCATGGGCCAAGAGTCTCTCCCTGCGG	37135
QY	781	ATCCGACCCCTCCCTGGGCCCATCTCAAGGCTGCCCCCTTCTCATCCTACTTGGGATTC	840

Db	37134	ATCCGACCCCTCCCTGGGGCCATCTCAAGGCTGCCCTTCTCTCACTACTTCGGACTC	37075
Qy	841	TTTCAGGTTCACTGAGGGGCCCTGTGTCGCCACAGTCGTCCAGGCTGCGGCTCCCT	900
Db	37074	TTTCAGGTTCACTGAGGGGCCCTGTGTCGCCAGTCGTCCAGGCTCCGGCTCCCT	37015
Qy	901	GCAGAGCTCTGGGGACCGCGGTCCCTCTGGCCCCAACCCCTCAGCGGCTTTGGTCCAGA	960
Db	37014	GCAGAGCTCTGGGGACCGCGGTCCCTCTGGCCCCAACCCCTCAGCGGCTTTGGTCCAGA	36955
Qy	961	CCGTGCCCTCCCTCAGAGGTCGCTGGGCGCTGTTCACGTGTTTTCATCCACATAAAT	1020
Db	36954	CCGTGCCCTCCCTCAGAGGTCGCTGGGCGCTGTTCACGTGTTTTCATCCACATAAAT	36895
Qy	1021	ACAGTATTCCTCACTCTTATCTTACAACTCCCCACCGCCACTCTCCACTCACTAGCTC	1080
Db	36894	ACAGTATTCCTCACTCTTATCTTACAACTCCCCACCGCCACTCTCCACTCACTAGCTC	36835
Qy	1081	CCCAATCCCTGACCCCTTTAGGGCCCCAGTAGTCTGACTTCCGCCCTGGCCAGAGCCCC	1140
Db	36834	CCCAATCCCTGACCCCTTTAGGGCCCCAGTAGTCTGACTTCCGCCCTGGCCAGAGCCCC	36775
Qy	1141	CAGGGCATTTGTCTCACTGACTCTGTGTGGGCAAGATGGGTCCAGAAAGCACCACCTTCAG	1200
Db	36774	CAGGGCATTTGTCTCACTGACTCTGTGTGGGCAAGATGGGTCCAGAAAGCACCACCTTCAG	36715
Qy	1201	GCACCTAAGGGGGCTGTGACCTGTGGCGCAGAAAGCCAAAGAGCTGGGCTTAGGCCAGAGG	1260
Db	36714	GCACCTAAGGGGGCTGTGACCTGTGGCGCAGAAAGCCAAAGAGCTGGGCTTAGGCCAGAGG	36655
Qy	1261	TTTCCCAAATGTAGGGGCGAGAAACAAAGCAAGCTCTCCCTGTGAAATTCCTGTGAT	1320
Db	36654	TTTCCCAAATGTAGGGGCGAGAAACAAAGCAAGCTCTCCCTGTGAAATTCCTGTGAT	36595
Qy	1321	TTTTAAACAGATATTTATTTATTTATTTATTTGTGTGACAAATGTTGATTAATGG	1373
Db	36594	TTTTAAACAGATATTTATTTATTTATTTATTTGTGTGACAAATGTTGATTAATGG	36542

RESULT 7  
 AC127470  
 LOCUS  
 DEFINITION  
 AC127470 215795 bp DNA linear HTG 17-JUL-2002  
 Pan troglodytes clone RP43-145D13, WORKING DRAFT SEQUENCE, 15  
 unordered pieces.  
 AC127470  
 AC127470.1 GI:2186866  
 HTG: HTGS\_PHASE1, HTGS\_DRAFT.  
 chimpanzee.  
 ORGANISM  
 Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.  
 REFERENCE  
 1 (bases 1 to 215795)  
 Akher,N., Antonellis,A., Ayele,K., Beckstrom-Stenberg,S.M.,  
 Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,  
 Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,  
 Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Latic,P.,  
 Lee,Li,S.-Q., Legaapi,R., Maduro,Q.L., Maduro,V.B.,  
 Marquies,E.H., Masiello,C., Maskerl,B., Mastriani,S.D.,  
 McCloskey,J.C., McDowell,J., Paguligan,C., Pearson,R.,  
 Portnoy,M.E., Prasad,A., Schueler,M.G., Stantirpop,S., Thomas,J.W.,  
 Thomas,P.J., Touchman,J.W., Tsurugano,C., Vogt,J.L., Walker,M.A.,  
 Welthery,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.  
 NISC Comparative Sequencing Initiative  
 UNPUBLISHED  
 2 (bases 1 to 215795)  
 Green,E.D.  
 Direct Submission  
 Submitted (17-JUL-2002) NIH Intramural Sequencing Center, 8717  
 Government Circle, Gaithersburg, MD 20877, USA  
 ----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 COMMENT  
 Center code: NISC







Db 182 TTGAAACAACTAGTCGGCCCTCGAAGAGTCTCTTAAGCCCGAAGGCCCGGCTTGC 241  
 QY 418 AGACGATCGAGCCCATTTATGAAATTCATCCAGACCTGGACAGGAGCGGAGCA 477  
 Db 242 CGAGCTATGACAGCCCATTTATGAGTTTATCTCTGGCCAGACAGATGAGACACAAACA 301  
 QY 478 GGTGTGACGGGACAGTGTGCTGGGAGGAGCAACCAATCAACAGCTCCAGCCCTTG 537  
 Db 302 GGTGTGATGGGACAGTGTGCTGGGAGGAGCAACCAATCAACAGCTCCAGCCCTTG 361  
 QY 538 CGGTACAAACCCGATGAGGAGGAGTTTATCTACACCCGCTGGGCTTACTACTGTAC 597  
 Db 362 CGGTACGACCCGATGAGGAGGAGTTTATCTACACGCTGGGCTTACTACTGTAC 421  
 QY 598 TGTGAGTGCATTTGATGAGGAGGAGCTCTACCTGAGAGCTGAGTCTGCTGAT 657  
 Db 422 TGTGAGTGCATTTGATGAGGAGGAGCTCTACCTGAGAGCTGAGTCTGCTGAT 481  
 QY 658 GGTGTGCTGGCCCTGCTGCTGGGAGGAGTCTACGACCTGCGGCAATTCCTGCGG 717  
 Db 482 GGTGTGCTGGCCCTGCTGCTGGGAGGAGTCTACGACCTGCGGCAATTCCTGCGG 541  
 QY 718 CCCAGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 777  
 Db 542 CCCAGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601  
 QY 778 CGGATCGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837  
 Db 602 CGGATCGGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 661  
 QY 838 CTCTTCACAGTTCTACGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 897  
 Db 662 CTCTTCACAGTTCTACGAGGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721  
 QY 898 -----CCGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 952  
 Db 722 AGGAGCATACACACCTCTTACCCACCCCACTCTGCAACCCCTGCTGCTGCTGCTG 780  
 QY 953 GCTCCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1012  
 Db 781 GGTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837  
 QY 1013 ACATTAATACAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1072  
 Db 838 -----ACAGACGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 891  
 QY 1073 ACTAGCTCCCAATCCCTGAGACCTTGTGAGGCCCAAGATGATGCTGCTGCTGCTGCTG 1132  
 Db 892 ACTAGCTCCCAATCCCTGAGACCTTGTGAGGCCCAAGATGATGCTGCTGCTGCTGCTG 936  
 QY 1133 CAGACCCCAAGGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1192  
 Db 937 CACCCGACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 968  
 QY 1193 CACTTCAGGAGCTAAGAGGAGGCTGAGACCTGCGGAGGAGCAACCAAGATGAGGCTGAG 1252  
 Db 969 -----CAGGACGATGAGGAGGCTGAGACCTGCGGAGGAGCAACCAAGATGAGGCTGAG 1023  
 QY 1253 GCCAGAGATGCTCAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1311  
 Db 1024 GCCAGAGATGCTCAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1079  
 QY 1312 CCTGTGATTTTAAACAGATATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 1366  
 Db 1080 CCTGTGATTTTAAACAGATATTTTATTTATTTATTTATTTATTTATTTATTTATTTA 1132  
 RESULT 10  
 AC126921  
 LOCUS AC126921 153553 bp DNA linear HTG 10-JUL-2002  
 DEFINITION Bos taurus clone RP42-45D24, WORKING DRAFT SEQUENCE, 12 unordered  
 pieces.

ACCESSION AC126921  
 VERSION AC126921.1 GI:21724098  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE  
 ORGANISM  
 Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.  
 REFERENCE  
 1 (bases 1 to 153553)  
 Ahlert, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,  
 Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,  
 Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Latic, P.,  
 Lee-Ilin, S.-Q., Legaspi, R., Maduro, Q.L., Maduro, V.B.,  
 Margulies, E.H., Mastello, C., Masker, B., Mastrian, S.D.,  
 McCloskey, J.C., McDowell, J., Paquirigan, C., Pearson, R.,  
 Portnoy, M.E., Prasad, A., Scheuler, M.G., Stantip, S., Thomas, J.W.,  
 Thomas, P.J., Touchman, J.W., Tsuneyon, C., Vogt, J.L., Walker, M.A.,  
 Werber, K.D., Wiggins, L., Young, A., Zhang, L., H. and Green, E.D.  
 NISC Comparative Sequencing Initiative  
 2 (bases 1 to 153553)  
 Unpublished  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 JOURNAL  
 TITLE  
 JOURNAL  
 COMMENT  
 Direct Submission  
 Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gaithersburg, MD 20877, USA  
 ----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: <http://www.nisc.nih.gov>  
 Contact: nisc-zoo@nigrl.nih.gov  
 Project Information  
 Center project name: ddi  
 Center clone name: 045D24  
 ----- Summary Statistics  
 Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 146066 bases at least Q40  
 Consensus quality: 147748 bases at least Q30  
 Consensus quality: 148824 bases at least Q20  
 Insert size: 151000; agarose-fp  
 Insert size: 152453; sum-of-ctrls  
 Quality coverage: 8.80x in Q20 bases; agarose-fp  
 Quality coverage: 8.72x in Q20 bases; sum-of-ctrls  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 12 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 1  
 2312: contig of 2312 bp in length  
 2313 2412: gap of unknown length  
 2413 5841: contig of 3429 bp in length  
 5842 5941: gap of unknown length  
 5942 8435: contig of 2494 bp in length  
 8436 8535: gap of unknown length  
 8536 15799: contig of 7264 bp in length  
 15800 15899: gap of unknown length  
 15900 25224: contig of 9325 bp in length  
 25224 25324: gap of unknown length  
 25325 32504: contig of 7180 bp in length  
 32505 32604: gap of unknown length  
 32605 40970: contig of 8366 bp in length  
 40971 41070: gap of unknown length  
 41071 56590: contig of 15520 bp in length  
 56591 56690: gap of unknown length  
 56691 73769: contig of 17079 bp in length  
 73770 73869: gap of unknown length  
 73870 90859: contig of 16990 bp in length



Center clone name: RP23-168P5  
 ----- Summary Statistics  
 Sequencing vector: M13; L08821  
 Chemistry: Dye-terminator Big Dye; 48% of reads  
 Assembly program: Phrap; version 0.990329  
 Consensus quality: 212648 bases at least Q40  
 Consensus quality: 218902 bases at least Q30  
 Consensus quality: 222384 bases at least Q20  
 Estimated insert size: 210556; sum-of-contigs estimation  
 Quality coverage: 0x in Q20 bases; agarose-gel estimation  
 Quality coverage: 7.2x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length  
 (see <http://www.hgsc.bcm.tmc.edu/docs/genbank.draft.data.html>).  
 NOTE: This is a 'working draft' sequence. It currently  
 consists of 7 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1 62152: contig of 62152 bp in length  
 \* 62153 62252: gap of unknown length  
 \* 62253 118772: contig of 56520 bp in length  
 \* 118773 118872: gap of unknown length  
 \* 118873 148924: contig of 30052 bp in length  
 \* 148925 149024: gap of unknown length  
 \* 149025 167231: contig of 18207 bp in length  
 \* 167232 167331: gap of unknown length  
 \* 167331 189907: contig of 22576 bp in length  
 \* 189908 190007: gap of unknown length  
 \* 190008 196537: contig of 6530 bp in length  
 \* 196538 196637: gap of unknown length  
 \* 196638 203083: contig of 6446 bp in length.

FEATURES  
 source  
 1. 203083  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /chromosome="11"  
 /clone="RP23-168P5"

BASE COUNT 52662 a 49293 c 47892 g 52633 t 603 others  
 ORIGIN

Query Match 22.5%; Score 309.6; DB 2; Length 203083;  
 Best Local Similarity 72.0%; Pred. No. 9.5e-47;  
 Matches 556; Conservative 0; Mismatches 149; Indels 67; Gaps 9;

QY 601 CAGGTGACCTTGATGAGGGGAGGAGGCTGCTCACTGAAAGCTGCACTTGTGATGATGAT 660  
 |||||  
 Db 43019 CAGGTGACCTTGATGAGGGGAGGAGGCTGCTCACTGAAAGCTGCACTTGTGATGATGAT 42960  
 QY 661 GTGTGGGCGCTGCGCTGCGCTGAGGAATTCACGCACTGCGGCGAGTTCCCTGGGGCC 720  
 |||||  
 Db 42959 GTGTGGGCGCTGCGCTGCGCTGAGGAATTCACGCACTGCGGCGAGTTCCCTGGGGCC 42900  
 QY 721 CAGTCCGCTCTGCAAGGCTGTGGGCTGTGGCCCTGCGGCGAGGATCTCCCTGGCG 780  
 |||||  
 Db 42899 CAGTCCGCTCTGCAAGGCTGTGGGCTGTGGCCCTGCGGCGAGGATCTCCCTGGCG 42840  
 QY 781 ATCCGCACTTCCCTGCGGCGCAATCTCAAGCTGCGGCGCTTCTCACTTCTGGAGTC 840  
 |||||  
 Db 42839 ATCCGCACTTCCCTGCGGCGCAATCTCAAGCTGCGGCGCTTCTCACTTCTGGAGTC 42780  
 QY 841 TTCAGGCTCACTGAGGGGCGCTGCTCCGCCACAGATGCTCCAGGCTGCGGCGTC--- 897  
 |||||  
 Db 42779 TTCAGGCTCACTGAGGGGCGCTGCTCCGCCACAGATGCTCCAGGCTGCGGCGTC--- 42720  
 QY 898 -CCTGACAGCTCTCTGGGCGAGCCGGTCCCTCTGCGGCCACCTCAAGCCGCTTTTGGT 955  
 |||||  
 Db 42719 AGCATGACAGCACTCTCCCTACCCACCCGCTCTCTCCACCCCTC-GCTGCTCTTGGT 42661

QY 956 CCAGACCTGCCCCCTCCCTCTGAGAGGCTGCGGGCGCTGTACGCTTTTCCATCCACCA 1015  
 |||||  
 Db 42660 CCAGACCTGCCCCCTCCCTCTGAGAGGCTGCGGGCGCTGTACGCTTTTCCATCCACCA 42602  
 QY 1016 TAAATACGATTTTCCCTCACTTATCTTACCTCCCGACGCCACCTCCACTACT 1075  
 |||||  
 Db 42601 GA-----CGTATCCCTGCTCTTCTTTTACATCCATCCACCACTATCCACTCAGT 42547  
 QY 1076 AGCTCCCAATCCCTGAGACCTTTGAGGCCCGCAGTATCTGACCTCCCGCTGGCCAG 1135  
 |||||  
 Db 42546 AGCTCCCAATCCCTGAGACCTTTGAGGCCCGCAGTATCTGACCTCCCGCTGGCCAG 42502  
 QY 1136 ACCCCAGGCGATTTGTTACTCTACTCTGTGTGGCAAGATGGGTCCAGAACCCAC 1195  
 |||||  
 Db 42501 CCGACCACTGTTATTTATCTTTGTGCAC----- 42473  
 QY 1196 TTCAGGCTCACTGAGGGGCTGAGCTGCGGCGAGAACCAAGACTGGCGCTAGGCC 1255  
 |||||  
 Db 42472 -CAGGCTCACTGAGGGGCTGAGCTGCGGCGAGAACCAAGACTGGCGCTAGGCC 42415  
 QY 1256 AGAGTTCCCAATGTTGAGGGGCGAGCA-MCAGACAAAGTCTCTCCCTTGAATTCCT 1314  
 |||||  
 Db 42414 AGAGTTCCCAATGTTGAGGGGCGAGCA-MCAGACAAAGTCTCTCCCTTGAATTCCT 42359  
 QY 1315 GTGATTTTAAACAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1366  
 |||||  
 Db 42358 GTGATTTTAAACAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 42309

RESULT 12  
 AL603707/c 234182 bp DNA linear ROD 17-NOV-2001  
 LOCUS  
 DEFINITION  
 Mouse DNA sequence from clone RP23-422L16 on chromosome 11,  
 complete sequence.  
 AL603707  
 AL603707.5 GI:17017790

KEYWORDS  
 SOURCE  
 ORGANISM  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Direct Submission  
 Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquerry@sanger.ac.uk  
 On Nov 20, 2001 this sequence version replaced GI:16605765.  
 During sequence assembly data is compared from overlapping clones.  
 Where differences are found these are annotated as variations  
 together with a note of the overlapping clone name. Note that the  
 variation annotation may not be found in the sequence submission  
 corresponding to the overlapping clone, as we submit sequences with  
 only a small overlap as described above.  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats: all regions were covered by at least  
 one plasmid subclone or more than one M13 subclone; and the  
 assembly was confirmed by restriction digest. The following  
 abbreviations are used to associate primary accession numbers given  
 in the feature table with their source databases: Em: EMBL; SW:  
 SWISSPROT; Tr: TrEMBL; Wp: WormPep; Information on the WORMPEP  
 database can be found at  
<http://www.sanger.ac.uk/projects/C.elegans/wormpep>  
 from the RPI-23 Mouse PAC Library  
 constructed by the group of Pletier de Jong.  
 For further details see <http://www.choil.org/bacpac/home.htm>  
 VECTOR: pBAC3.6  
 This sequence is the entire insert of clone RP23-422L16.  
 Location/Qualifiers  
 1. 234182











Db	42211	CTCAGAGTCGCCCCACCCCGCTGACCTGCTGCGCCCTCTGACAGGCTGCTGGACTTG	42152
Qy	961	CTGCCCCCTCCCTAGAGGCTGCCCTGGGCGCTGTTCACGTGTGTTTCATCCCATTAAT	1020
Db	42151	TTCACTGTCTTCGGGACACATACATATTCCTCTCCCTCTGACACCCCATCT	42096
Qy	1021	ACAGTATCCCACTCTTATCTTACACATGCCCCACCGCCACTCTCACCTCAGCTC	1080
Db	42095	-----CCACGTCACCTGGGCGCCCTGAACCCCTGCTTTTGAAGCCCCATTTATCTC	42045
Qy	1081	CCCAATCCCTGACCCCTTTGAGGCGCCCACTGATCTGCACCTCCCGCTGGCCACAGACCC	1140
Db	42044	CTGACTTCTTAATCTTGTGGCAACCCCGCCACCCCGCTCCCGCCCAACACACAGC----	41989
Qy	1141	CAGGCATTGTGTTACTGTACTCTGTGGGCAAGATGGGTTCACAA---GACCCCACTT	1197
Db	41988	-----CAAAAGCACATGCTGTCTATCTGTGGGTGAGATGGGTCTCCTAACCTCCCGCCACTT	41934
Qy	1198	CAGGCATCAAGGGGGCTGAGACCTGGCGGCGAGGAAGCCAAAGACAGCTGGGCTAAGCCAG	1257
Db	41933	CAGGCATCAATCCAGGGCTAGAGCCAGGGCGCACTGGACAG-----GAAGCCGG	41884
Qy	1258	GAGTCCCAAAATGTAGGGGGGAGAAACAAAGCAAGCTCTCCCTTGAAGAAATCCCTGTG	1317
Db	41883	GAGTCCCAATCGCGCAGGGCTGAG-AACAGAGTGAAGCTC-----TTCCCTGTG	41836
Qy	1318	GATTTTAAAAACAGATATTTATTTTATTTATTTATTTGACAAAAATGTTGATAAATGG	1373
Db	41835	GATTTTAAAAACAGATATTTATTTTATTTATTTATTTGTCACAGATCTCAATTAAGTGG	41780

Search completed: May 8, 2003, 10:02:32  
Job time : 4943.14 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 11:47:25 ; Search time 1852 Seconds  
(without alignments)  
12006.697 Million cell updates/sec

Title: US-09-245-198a-3

Perfect score: 1373  
Sequence: 1 atgcattgttagacttga.....gacaaatgttataaatg 1373

Scoring table:  
OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size: 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vit:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	629	45.8	777	13	BI819200 603034614
2	613	44.6	1071	14	BM921213 AGENCOURT
3	605	44.1	963	14	BM921213 AGENCOURT
4	518	37.7	828	13	BI596681 603243254
5	506	36.9	728	13	BI870393 603395641
6	506	36.9	731	13	BI871711 603395625

Result No.	Score	Query Match	Length	DB ID	Description
7	491	35.8	609	13	BI966060 1e72904.x
8	437	31.8	948	14	BM707185 AGENCOURT
9	436	31.8	940	14	BM984231 AGENCOURT
10	416	30.3	834	13	BI766766 603056866
11	394	28.7	413	9	AI422796
12	394	28.7	568	14	BM971606 BI-CF-ECL1
13	375	27.3	531	13	BI824443 603038693
14	352	25.6	440	13	BM128059
15	349	25.4	951	14	BM674188 AGENCOURT
16	341	24.8	377	14	BM509016
17	329	24.0	710	12	BE858778
18	328	23.9	367	13	BM662622
19	327	23.8	346	12	BE858822
20	326	23.7	337	10	AW195034
21	326	23.7	345	12	BF439893
22	326	23.7	352	10	AM204512
23	326	23.7	364	10	AM291620
24	324	23.6	399	9	AI913541
25	324	23.6	456	13	BI966255
26	321	23.4	367	17	AQ100365
27	321	23.4	1819	14	BM925491
28	320	23.3	374	13	BM505649
29	318	23.2	329	9	AI695776
30	318	23.2	346	14	BM703512
31	316	23.0	407	12	BF195436
32	309	22.5	416	9	AI221985
33	307	22.4	910	12	BG110063
34	303	22.1	317	9	AI202121
35	298	21.7	346	14	BM703512
36	290	21.1	298	14	BM688946
37	283	20.6	587	12	BG686319
38	271	19.7	1027	17	AF163779
39	267	19.4	346	9	AI760777
40	262	19.1	346	12	BF940141
41	247	18.0	561	12	BG054878
42	236	17.2	318	10	AM081731
43	230	16.8	264	10	AM661741
44	229	16.7	253	9	AI863563
45	227	16.5	238	9	AI682487

## ALIGNMENTS

RESULT 1  
LOCUS: BI819200 777 bp mRNA linear EST 04-OCT-2001  
DEFINITION: 603034614F1 NIH\_MGC\_115 Homo sapiens cdna clone IMAGE:5175698 5',  
ACCESSION: BI819200  
KEYWORDS: BI819200.1 GI:15930750  
SOURCE: EST.  
ORGANISM: human.  
REFERENCE: NIH-MGC <http://mhc.nci.nih.gov/>.  
AUTHORS: National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE: Unpublished (1999)  
JOURNAL: Contact: Robert Strausberg, Ph.D.  
COMMENT: Email: [cgapbs-ri@mail.nih.gov](mailto:cgapbs-ri@mail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: <http://image.llnl.gov>  
Plate: LLM11437 row: 1 column: 03  
High quality sequence stop: 759.  
Location/Qualifiers: 1..777

FEATURES  
source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5175698"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPORT6; Site:1; NotI; Site:2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT      140 a      268 c      215 g      154 t
ORIGIN

```

```

Query Match      45.8%; Score 629; DB 13; Length 777;
Best Local Similarity 99.7%; Pred. No. 4.8e-313;
Matches 729; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 288 GGCAGAGAGACGACGACGACCGCTGGAACTGATCCGACAGAGAAAGCCAGATCC 347
    |||||||
Db 13 GGCAGAGAGACGACGACGACCGCTGGAACTGATCCGACAGAGAAAGCCAGATCC 72
QY 348 TGGCGCTTCTGAAACGACACTAGTTCGGCTTCGCGAAAGTGCACCTAAAGCCGGAATC 407
    |||||||
Db 73 TGGCGCTTCTGAAACGACACTAGTTCGGCTTCGCGAAAGTGCACCTAAAGCCGGAATC 132
QY 408 ACGGCTCGAAGAGCGATGCGACGCCATTATGAACTTCATCCAGCACTGGACAGGAGG 467
    |||||||
Db 133 ACGGCTCGAAGAGCGATGCGACGCCATTATGAACTTCATCCAGCACTGGACAGGAGG 192
QY 468 AGCGAGGACAGGTGAGACGAGGACAGTGTGCTGGAGAGAGGACAGATCAACAGCTC 527
    |||||||
Db 193 AGCGAGGACAGGTGAGACGAGGACAGTGTGCTGGAGAGAGGACAGATCAACAGCTC 252
QY 528 CAGGCTGTGGCTACAAACCGCCAGATGCGGAGTTTATGACACCGGCTGGCTCTTA 587
    |||||||
Db 253 CAGGCTGTGGCTACAAACCGCCAGATGCGGAGTTTATGACACCGGCTGGCTCTTA 312
QY 588 CTACCTGTACTGTAGGTGACACTTGTATGAGGAGGAGCTGTCTACCTGAAGCTGGAAT 647
    |||||||
Db 313 CTACCTGTACTGTAGGTGACACTTGTATGAGGAGGAGCTGTCTACCTGAAGCTGGAAT 372
QY 648 GCTGTGATGTGTGCTGCGCTGCGCTGCGTGGAGAAATTCAGGCACTGCGGCGAG 707
    |||||||
Db 373 GCTGTGATGTGTGCTGCGCTGCGCTGCGTGGAGAAATTCAGGCACTGCGGCGAG 432
QY 708 TTCCCTGGGGGCGGAGCGCCGCTGTGCAAGTGTCTGGGCTGTGGGCGCTGGCGGAG 767
    |||||||
Db 433 TTCCCTGGGGGCGGAGCGCCGCTGTGCAAGTGTCTGGGCTGTGGGCGCTGGCGGAG 492
QY 768 GTCTCTCCCTGCGGATCGGACACCTCCCTGGGCGCATCTCAAGGCTGCCCCCTTCCTAC 827
    |||||||
Db 493 GTCTCTCCCTGCGGATCGGACACCTCCCTGGGCGCATCTCAAGGCTGCCCCCTTCCTAC 552
QY 828 CTATTGGGACTCTTCAAGGTTCACTAGAGGGGCGCTGCTCCACAGTGTCCAGGC 887
    |||||||
Db 553 CTATTGGGACTCTTCAAGGTTCACTAGAGGGGCGCTGCTCCACAGTGTCCAGGC 612
QY 888 TGGCGGCGCCCTGAGACGCTCTGGGCAACCGGTCGCCCTCTGCCCCCAGCCTCAGCCG 947
    |||||||
Db 613 TGGCGGCGCCCTGAGACGCTCTGGGCAACCGGTCGCCCTCTGCCCCCAGCCTCAGCCG 672
QY 948 TCTTGTCTCAGACCTGCCCCCTCCTCTAGAGGCTGCTGGGCGCTGTTACAGTCTTTC 1007
    |||||||
Db 673 TCTTGTCTCAGACCTGCCCCCTCCTCTAGAGGCTGCTGGGCGCTGTTACAGTCTTTC 732
QY 1008 ATCCGACATTA 1018
    |||||||
Db 733 ATCCGACATTA 743

```

```

RESULT 2
BM921213
LOCUS
DEFINITION
BM921213
1071 bp mRNA linear EST 12-MAR-2002
AGENCOURT_6633046 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:51752561
5', mRNA sequence.
ACCESSION
BM921213
VERSION
BM921213.1 GI:19371592
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1071)
NIH-MGC http://mgc.ncl.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLM12786 row: P column: 02
High quality sequence stop: 656.

```

## FEATURES

source

1. 1071

location/Qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:51752561"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
PCMV-SPORT6; Site:1; NotI; Site:2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dr primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."
BASE COUNT      226 a      346 c      279 g      220 t
ORIGIN

```

```

Query Match      44.6%; Score 613; DB 14; Length 1071;
Best Local Similarity 99.7%; Pred. No. 8.1e-305;
Matches 713; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 412 GCTCGAAGAGGATGCGACGCCATTATGAACTTATCAGCACTGAGACAGGAGCG 471
    |||||||
Db 12 GCTCGAAGAGGATGCGACGCCATTATGAACTTATCAGCACTGAGACAGGAGCG 71
QY 472 CAGGACAGTGTGAGACGAGTGTGCTGGAGGAGAGCCAGAAATCAACAGCTCAGC 531
    |||||||
Db 72 CAGGACAGTGTGAGACGAGTGTGCTGGAGGAGAGCCAGAAATCAACAGCTCAGC 131
QY 532 CTTCTGCGCTACAAACCGCCAGATGCGGAGTTTATAGTCAACCGGCTGGGCTCTACTAC 591
    |||||||
Db 132 CTTCTGCGCTACAAACCGCCAGATGCGGAGTTTATAGTCAACCGGCTGGGCTCTACTAC 191
QY 592 CTGTACTGTAGGTGCACTTGTATGAGGAGGAGGAGCTCTTCACTGAGAGCTGAGTCTG 651
    |||||||
Db 192 CTGTACTGTAGGTGCACTTGTATGAGGAGGAGGAGCTTCACTGAGAGCTGAGTCTG 251
QY 652 GTGAGTGTGTGTGCGCTGCGCTGCGCTGAGGAATTCACGCACTGCGGCGAGTTCC 711
    |||||||
Db 252 GTGAGTGTGTGTGCGCTGCGCTGCGCTGAGGAATTCACGCACTGCGGCGAGTTCC 311

```

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (INTL.)

DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: L1AM1722 row: k column: 13  
 High quality sequence stop: 776.

## FEATURES

source

1. 828  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5285892"  
 /clone.lib="NIH\_MGC.96"  
 /tissue\_type="hypothalamus"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pBluescript (modified  
 pBluescript KS+); Site\_1: BamHI; Site\_2: SalI-XhoI (gtcgag  
 ); Oligo-dt primed using primer 5'-TTTTTTTTTTTTTTV-3',  
 size-selected for average insert size 2.3 kb and  
 normalized to 10<sup>5</sup>. This is a primary library enriched  
 for full-length clones and constructed using the  
 Cap-trapper method (Carninci, in preparation). Library  
 constructed by M. Brownstein (NIH/NHGRI, National  
 Institutes of Health). Note: this is a NIH\_MGC Library."  
 BASE COUNT 155 a 278 c 223 g 172 t  
 ORIGIN

## Query Match

Best Local Similarity 99.7%; Score 518; DB 13; Length 828;  
 Pred. No. 8.2e-256;

Matches 618; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

312 GGACGTAAATCCCGACAGAGAAAGCCAGATCCCGCTTCTCGAACCAGTACT 371  
 43 GGAAGTAAATCCCGACAGAGAAAGCCAGATCCCGCTTCTCGAACCAGTACT 102  
 372 TCGGCTCGAGAGATCCAGTAAAGCCGGAACACAGGCTCGAAGAGGATCGAGC 431  
 103 TCGGCTCGAGAGATCCAGTAAAGCCGGAACACAGGCTCGAAGAGGATCGAGC 162  
 432 CCATTATGAAGTATCCAGAGCTGAGACAGAGAGAGAGAGAGAGAGAGAGAG 491  
 163 CCATTATGAAGTATCCAGAGCTGAGACAGAGAGAGAGAGAGAGAGAGAGAG 222  
 492 AGTAGTGGCTGGAG 551  
 223 ATTGAGTGGCTGGAG 282  
 552 GATCGGGAGTTATAGTCAACCGGGCTGGCTTACTACTGATGATGATGATGAT 611  
 283 GATCGGGAGTTATAGTCAACCGGGCTGGCTTACTACTGATGATGATGATGAT 342  
 612 TGATGAGGAGAGAGCTGTCTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 671  
 343 TGATGAGGAGAGAGCTGTCTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 402  
 672 GCGCTGCTGGAGAGATTCACGACCTAGGAGAGAGAGAGAGAGAGAGAGAGAG 731  
 403 GCGCTGCTGGAGAGATTCACGACCTAGGAGAGAGAGAGAGAGAGAGAGAGAG 462  
 732 CTGCGAGGTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGT 791  
 463 CTGCGAGGTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGT 522  
 792 CCGCTGGGAGAGATTCACGACCTAGGAGAGAGAGAGAGAGAGAGAGAGAG 851  
 523 CCGCTGGGAGAGATTCACGACCTAGGAGAGAGAGAGAGAGAGAGAGAGAG 582  
 852 CTGAGGAGAGAGAGATTCACGACCTAGGAGAGAGAGAGAGAGAGAGAGAGAG 911  
 583 CTGAGGAGAGAGAGATTCACGACCTAGGAGAGAGAGAGAGAGAGAGAGAGAG 642  
 912 TGGGAGAGAGAGAGATTCACGACCTAGGAGAGAGAGAGAGAGAGAGAGAG 931  
 643 TGGGAGAGAGAGAGATTCACGACCTAGGAGAGAGAGAGAGAGAGAGAGAG 662

RESULT 5  
 BI870393  
 LOCUS  
 DEFINITION 60339564.F1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:5405459 5',  
 mRNA sequence.

ACCESSION BI870393  
 VERSION BI870393.1 GI:16044066  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 728)

NIH-MGC <http://imgc.ncl.nih.gov/>

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LNL at:

<http://image.llnl.gov>

Plate: L1AM12034 row: a column: 12

High quality sequence stop: 728.

Location/Qualifiers

## FEATURES

source

1. 728  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5405459"  
 /clone.lib="NIH\_MGC.90"  
 /tissue\_type="adenoecarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: liver; Vector: pCMV-Sport6; Site\_1: NotI;  
 Site\_2: SalI; Cloned unidirectionally; Oligo-dt primed.  
 Average insert size 1.7 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC library."  
 BASE COUNT 125 a 240 c 227 g 136 t  
 ORIGIN

## Query Match

Best Local Similarity 100.0%; Score 506; DB 13; Length 728;  
 Pred. No. 1.3e-249;

Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

262 CAGGAGCTGCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 321  
 74 CAGGAGCTGCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 133  
 322 CCCGAGACAG 381  
 134 CCCGAGACAG 193  
 382 AGAAGTCACTAAAGCCGGAACACAGGAGCTGAGAGAGAGAGAGAGAGAGAGAG 441  
 194 AGAAGTCACTAAAGCCGGAACACAGGAGCTGAGAGAGAGAGAGAGAGAGAGAG 253  
 442 GTTCATCCAG 501  
 254 GTTCATCCAG 313  
 502 TGGGAG 561  
 314 TGGGAG 373  
 562 TTTATAGTCAACCGGGCTGGCTTACTACTTACTAGTCAAGTCAAGTCAAGTCAAG 621  
 374 TTTATAGTCAACCGGGCTGGCTTACTACTTACTAGTCAAGTCAAGTCAAGTCAAG 433

```

OY 622 AAGGCTGTACCTGAAGCTGACCTGTGTGTGATGTGTGTGCTGGCCCTGGCCCTGCTG 681
    |||||||
DB 434 AAGGCTGTACCTGAAGCTGACCTGTGTGTGATGTGTGTGCTGGCCCTGGCCCTGCTG 493
OY 682 GAGGAATTCAGACCACTGCGGCGCACTTCGCGGCCCCAGCTCCGCTTGCAGGTG 741
    |||||||
DB 494 GAGGAATTCAGACCACTGCGGCGCACTTCGCGGCCCCAGCTCCGCTTGCAGGTG 553
OY 742 TCTGGGCTGTGGCCCTGGCCCTGAGG 767
    |||||||
DB 554 TCTGGGCTGTGGCCCTGGCCCTGAGG 579

RESULT 6
BI871711 731 bp mRNA linear EST 11-OCT-2001
LOCUS 603395825F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405478 5',
DEFINITION mRNA sequence.
ACCESSION BI871711
VERSION BI871711.1 GI:16045386
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/
AUTHORS 1 (bases 1 to 731)
TITLE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-riemail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/ILNI at:
http://image.llnl.gov
Plate: LHAM12034 row: b column: 07
High quality sequence stop: 728.
Location/Qualifiers
1. 731
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5405478"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: Liver; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH-MGC Library."

BASE COUNT 125 a 242 c 228 g 136 t
ORIGIN

Query Match 36.9%; Score 506; DB 13; Length 731;
Best Local Similarity 100.0%; Pred. No. 1.3e-249;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 262 CAGGAGCCTGCCAGAGAGCTGTGTGTGAGAGAGACAGACCCGCGGAACGAGT 321
    |||||||
DB 74 CAGGAGCCTGCCAGAGAGCTGTGTGTGAGAGAGACAGACCCGCGGAACGAGT 133
OY 322 CCCAGACAGAAAGAACCCAGAGATCTGTGCGCTTTCCTGAACGAGTATTTGGCCTGCG 381
    |||||||
DB 134 CCCAGACAGAAAGAACCCAGAGATCTGTGCGCTTTCCTGAACGAGTATTTGGCCTGCG 193
OY 382 AAGATGACCTTAAAGGCGGGAAGACAGCGGCTCGAAGCGATGCGACCATTAATGAA 441
    |||||||
DB 194 AAGATGACCTTAAAGGCGGGAAGACAGCGGCTCGAAGCGATGCGACCATTAATGAA 253
OY 442 GTTCATCCAGACTGAGACGAGGCGAGGCGAGGTGTGAGCGGAGCATGAGTGGC 501

```

```

DB 254 GTTCATCCAGACTGAGACGAGAGCGAGCGAGGCTGTGAGCGGACATGAGTGGC 313
    |||||||
OY 502 TGGGAGGAGCCAGAAATCAACAGCTTCAGCCCTTGCGGTACAAACCCAGATGGGAG 561
    |||||||
DB 314 TGGGAGGAGCCAGAAATCAACAGCTTCAGCCCTTGCGGTACAAACCCAGATGGGAG 373
OY 562 TTTATAGTACCCGGGCTGGGCTCTACTACCTGTACTGTACAGTGCACTTATAGAGGG 621
    |||||||
DB 374 TTTATAGTACCCGGGCTGGGCTCTACTACCTGTACTGTACAGTGCACTTATAGAGGG 433
OY 622 AAGGCTGTACCTGAAGCTGACCTGTGTGTGATGTGTGTGCTGGCCCTGGCCCTGCTG 681
    |||||||
DB 434 AAGGCTGTACCTGAAGCTGACCTGTGTGTGATGTGTGTGCTGGCCCTGGCCCTGCTG 493
OY 682 GAGGAATTCAGACCACTGCGGCGCACTTCGCGGCCCCAGCTCCGCTTGCAGGTG 741
    |||||||
DB 494 GAGGAATTCAGACCACTGCGGCGCACTTCGCGGCCCCAGCTCCGCTTGCAGGTG 553
OY 742 TCTGGGCTGTGGCCCTGGCCCTGAGG 767
    |||||||
DB 554 TCTGGGCTGTGGCCCTGGCCCTGAGG 579

RESULT 7
BI966060 609 bp mRNA linear EST 12-MAR-2002
LOCUS le72904.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
DEFINITION cDNA clone IMAGE:5672623 3' similar to TR:054907 054907 TNF-RELATED
WEAK INDUCER OF APOPTOSIS ;, mRNA sequence.
ACCESSION BI966060
VERSION BI966060.1 GI:16340465
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/
AUTHORS 1 (bases 1 to 609)
TITLE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brownjfas.harvard.edu) This sequence now available from the IMAGE
consortium, for clone orders contact: info@image.llnl.gov
High quality sequence stop: 412.
Location/Qualifiers
1. 609
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5672623"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/notes="Organ: Pancreas; Vector: pSPORT1; Site.1: Not 1;
Site.2: Sal 1; Starting library constructed using
Superscript Plasmid Library kit (Life Technologies). cDNA

```

made by oligo-dt priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an EcoI of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 142 a 128 c 209 g 129 t 1 others  
ORIGIN

Query Match 35.8%; Score 491; DB 13; Length 609;  
Best Local Similarity 100.0%; Pred. No. 6.9e-242;  
Matches 491; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 883 CAGCTGCGGCTCCCTCGACAGCTCTGTGGACACGGTCCCTCGCCACCTCA 942  
|||||  
DB 513 CAGCTGCGGCTCCCTCGACAGCTCTGTGGACACGGTCCCTCGCCACCTCA 454  
943 GCGCTCTTTGGCTCGACAGCTGCCCCCTCTAGAGGCTGCGGCTGTTACGTC 1002  
DB 453 GCGCTCTTTGGCTCGACAGCTGCCCCCTCTAGAGGCTGCGGCTGTTACGTC 394  
OY 1003 TTTCCATCCCATATAAATACAGTATCCGACCTTATCTTACACTCCCCACGCCAC 1062  
|||||  
DB 393 TTTCCATCCCATATAAATACAGTATCCGACCTTATCTTACACTCCCCACGCCAC 334  
OY 1063 TCTCCACTCACTAGCTCCCAATCCCTGACCTTTGAGGCCCCCAGATGTGACATC 1122  
|||||  
DB 333 TCTCCACTCACTAGCTCCCAATCCCTGACCTTTGAGGCCCCCAGATGTGACATC 274  
OY 1123 CCCCTGGCCACAGACCCCGGAGGCAATGTGTTACTGCTGTGGCAAGATGGGTC 1182  
|||||  
DB 273 CCCCTGGCCACAGACCCCGGAGGCAATGTGTTACTGCTGTGGCAAGATGGGTC 214  
OY 1183 CAGAGACCCCATCTCAGGCACTAGAGGGCTGAGGCTGAGGCGGAGAGCAAGAGA 1242  
|||||  
DB 213 CAGAGACCCCATCTCAGGCACTAGAGGGCTGAGGCTGAGGCGGAGAGCAAGAGA 154  
OY 1243 CTGGGCTTAGGCGCAGAGATTCCTCAATGTGAGGGCGAGAAACAAGACTCTCCT 1302  
|||||  
DB 153 CTGGGCTTAGGCGCAGAGATTCCTCAATGTGAGGGCGAGAAACAAGACTCTCCT 94  
OY 1303 TGAGATTCCTCTGTGGATTTTAAACAGATATTTATTTATTTATTTGACAAATG 1362  
|||||  
DB 93 TGAGATTCCTCTGTGGATTTTAAACAGATATTTATTTATTTATTTGACAAATG 34  
OY 1363 TTGATTAATGG 1373  
|||||  
DB 33 TTGATTAATGG 23

RESULT 8 948 bp mRNA linear EST 16-JUL-2002  
B0707185  
LOCUS  
DEFINITION AGENCOURT\_8353983 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:6278608  
5', mRNA sequence.

ACCESSION B0707185  
VERSION B0707185.1 GI:21846084  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheraia; Primates; Catarrhini; Homidae; Homo.  
1 (bases 1 to 948)  
NIH-MGC http://mgi.ncl.nih.gov/.

AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
TITLE Unpublished (1999)  
JOURNAL Contact: Robert Strausberg, Ph.D.  
COMMENT Email: cgaaps-remail.nih.gov

Tissue Procurement: Dr. Mark Watson  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMU at:  
http://image.liml.gov  
Plate: LCM2466 row: n column: 17  
High quality sequence start: 24  
High quality sequence stop: 550.

FEATURES  
SOURCE

location/qualifiers

1. 948  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6278608"  
/clone\_11b="NIH\_MGC\_113"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: spleen; Vector: pORF7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."  
BASE COUNT 173 a 334 c 252 g 187 t 2 others  
ORIGIN

Query Match 31.8%; Score 437; DB 14; Length 948;  
Best Local Similarity 99.8%; Pred. No. 5.1e-214;  
Matches 487; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 475 GCAGGTGTGACGGGACAGTGTGCTGGAGAGCAAGCAATCAACAGCTCCAGCCCT 534  
|||||  
DB 167 GCAGGTGTGACGGGACAGTGTGCTGGAGAGCAAGCAATCAACAGCTCCAGCCCT 226  
OY 535 CTGGCTCAACACCGCAATGTGGAGATTATATGACACCGGCTGGGCTTACTACTG 594  
|||||  
DB 227 CTGGCTCAACACCGCAATGTGGAGATTATATGACACCGGCTGGGCTTACTACTG 286  
OY 595 TACTGTCAAGTCACTTGTATGATGAGGAGAGCTGTCTGATGAGTGAAGTGTGCTG 654  
|||||  
DB 287 TACTGTCAAGTCACTTGTATGATGAGGAGAGCTGTCTGATGAGTGAAGTGTGCTG 346  
OY 655 GATGTGTGTGCGGCTGCTGCTGTGAGGAAATTCACGACTGCGGCAAGTTCCTC 714  
|||||  
DB 347 GATGTGTGTGCGGCTGCTGCTGTGAGGAAATTCACGACTGCGGCAAGTTCCTC 406  
OY 715 GGGGCCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 774  
|||||  
DB 407 GGGGCCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 466  
OY 775 CTGGGATCCGACCTCCCTGCGGCAATCAAGCTGCGGCTGCTGCTGCTGCTGCTGCTG 834  
|||||  
DB 467 CTGGGATCCGACCTCCCTGCGGCAATCAAGCTGCGGCTGCTGCTGCTGCTGCTGCTG 526  
OY 835 GCATCTTCCAGGTTCACTAGAGGCGCTGTCCACAGTGTGCCAGGCTGCCGCGC 894  
|||||  
DB 527 GCATCTTCCAGGTTCACTAGAGGCGCTGTCCACAGTGTGCCAGGCTGCCGCGC 586  
OY 895 TCCCTGACAGCTCTGCTGCGGACCGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 954  
|||||  
DB 587 TCCCTGACAGCTCTGCTGCGGACCGGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTG 646  
OY 955 TCCAGACC 962  
|||||  
DB 647 TCCAGACC 654

RESULT 9 940 bp mRNA linear EST 16-AUG-2002  
B0884231  
LOCUS  
DEFINITION AGENCOURT\_8682031 Lupekl\_sciatic\_nerve Homo sapiens cDNA clone



IMAGE:6197488 5' mRNA sequence.  
ACCESSION B0884231  
VERSION B0884231.1 GI:2276239  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 940)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Dr. James R. Lupski  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: L1AM13607 row: j column: 17  
High quality sequence stop: 453.  
Location/Qualifiers  
1. 940  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="6197488"  
/clone\_lib="Lupski\_sciatic\_nerve"  
/sex="male"  
/tissue\_type="sciatic nerve"  
/dev\_stage="adult, 70 yr"  
/lab\_host="DH10B"  
/note="Vector: pCMV-SPORT6 (Life Technologies); Site\_1:  
Motl; Site\_2: SalI; cDNA made by oligo-dT priming.  
Directionally cloned using the following adaptors:  
5'-TCGACCCACGCGTCG-3' and  
5'-GACTAGTCTAGATCGAGCGCGCCCTT(15)-3'. Size selected >  
1 kb for average insert length 1.87 kb. This is a primary  
library, non-amplified. Library constructed by Life  
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor  
College of Medicine) and is available through Life  
Technologies."

BASE COUNT 170 a 325 c 255 g 190 t  
ORIGIN  
Query Match 31.8%; Score 436; DB 14; Length 940;  
Best Local Similarity 100.0%; Pred. No. 1.7e-213;  
Matches 436; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 301 CAGGACCGCTGGAGACTAAATCCAGACAGAAAGCCAGATCTGGCCTTCTG 360  
DB 1 CAGGACCGCTGGAGACTAAATCCAGACAGAAAGCCAGATCTGGCCTTCTG 60  
OY 361 AACGCTAGTTCGGCTCGCAGAAATGACCTAAAGCCCGGAAACAGCGGCTCGAAGA 420  
DB 61 AACGCTAGTTCGGCTCGCAGAAATGACCTAAAGCCCGGAAACAGCGGCTCGAAGA 120  
OY 421 GCGATCGACGCCATTTAAGAGTTCATCCAGACCTGGAGACGAGCAGCAGCAGGT 480  
DB 121 GCGATCGACGCCATTTAAGAGTTCATCCAGACCTGGAGACGAGCAGCAGCAGGT 180  
OY 481 GTGAGCGGAGACGTAGTGGCTGGAGAGGACCAATCAACAGTCCAGCCCTCTGGCC 540  
DB 181 GTGAGCGGAGACGTAGTGGCTGGAGAGGACCAATCAACAGTCCAGCCCTCTGGCC 240  
OY 541 TACAACCCGACATGGGAGATTATAGTCAACCCGGGCTGGGCTACTACTACTACTGT 600  
DB 241 TACAACCCGACATGGGAGATTATAGTCAACCCGGGCTGGGCTACTACTACTACTGT 300  
OY 601 CAGGTGACATTGATGAGAGGAGGCTGTCTACTGGAAGCTGACTTGTGATGATGT 660  
DB 301 CAGGTGACATTGATGAGAGGAGGCTGTCTACTGGAAGCTGACTTGTGATGATGT 360

661 GTGCTGCCCTGCTGCTGGAGAAATTCAGCACTGCGGCGCAGTTCCTCGGGCC 720  
DB 361 GTGCTGCCCTGCTGCTGGAGAAATTCAGCACTGCGGCGCAGTTCCTCGGGCC 420  
OY 721 CAGTCCGCTCTGCGC 736  
DB 421 CAGTCCGCTCTGCGC 436  
RESULT 10  
BI766766 834 bp mRNA linear EST 25-SEP-2001  
LOCUS 60305686P1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5206217 5',  
DEFINITION mRNA sequence.  
ACCESSION BI766766  
VERSION BI766766.1 GI:15758344  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 834)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
Plate: L1AM1517 row: c column: 18  
High quality sequence stop: 772.  
Location/Qualifiers  
1. 834  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_image="5206217"  
/clone\_lib="NIH\_MGC\_122"  
/lab\_host="DH10B"  
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;  
Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source  
anonymous pool of 24 week female lung, 16 week female  
spleen, and 20-22 week male spleens. Library is oligo-dT  
primed and directionally cloned (EcoRV site is destroyed  
upon cloning). Average insert size 1.4 kb. Insert size  
range 1-3 kb. Library is normalized and enriched for  
full-length clones and was constructed by C. Gruber  
(Invitrogen). Research Genetics tracking code 026. Note:  
this is a NIH-MGC Library."

BASE COUNT 154 a 287 c 223 g 170 t  
ORIGIN  
Query Match 30.3%; Score 416; DB 13; Length 834;  
Best Local Similarity 99.7%; Pred. No. 3.4e-203;  
Matches 656; Conservative 0; Mismatches 0; Indels 2; Gaps 2;  
OY 272 CCCAGAGAGCTGTGGAGAGGAGCAGGACCCGTCGGAATGATCCCGACAG 331  
DB 1 CCCAGAGAGCTGTGGAGAGGAGCAGGACCCGTCGGAATGATCCCGACAG 60  
OY 332 AAGAAAGCCAGATCCTGGCTTCCTGAACGCACTAGTTCGGCTCGCAGAACTGCAC 391  
DB 61 AAGAAAGCCAGATCCTGGCTTCCTGAACGCACTAGTTCGGCTCGCAGAACTGCAC 120  
OY 392 CTAAGGCGCGGAAACAGCGGCTCGAAGAGCATGCGACCCATTTATGAATTCTATCCAC 451  
DB 121 CTAAGGCGCGGAAACAGCGGCTCGAAGAGCATGCGACCCATTTATGAATTCTATCCAC 180







synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 165 a 286 c 315 g 184 t 1 others

ORIGIN

Query Match 25.4%; Score 349; DB 14; Length 951;

Best Local Similarity 100.0%; Pred. No. 1.3e-168;

Matches 349; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 369 AGTTGGGCTCGCAGAAAGTCACCTAAAGCCGGAAACACGGGCTCGAAGAGCGATCGC 428
    |||||||
Db 1 AGTTGGGCTCGCAGAAAGTCACCTAAAGCCGGAAACACGGGCTCGAAGAGCGATCGC 60
    |||||||

QY 429 AGCCATTATGAAGTTATCATCCAGACCTGGACAGAGGAGCGGAGGCTGGAACGG 488
    |||||||
Db 61 AGCCATTATGAAGTTATCATCCAGACCTGGACAGAGGAGCGGAGGCTGGAACGG 120
    |||||||

QY 489 GACAGTAGTGGCTGGGAGGAGGACAGATCAACAGCTCCAGCCCTGCGCTACAAACGG 548
    |||||||
Db 121 GACAGTAGTGGCTGGGAGGAGGACAGATCAACAGCTCCAGCCCTGCGCTACAAACGG 180
    |||||||

QY 549 CCAGATGGGGAGCTTTATAGTCAACCCGGGCTGGGCTCTACTACTGTACTGTCAAGTGCA 608
    |||||||
Db 181 CCAGATGGGGAGCTTTATAGTCAACCCGGGCTGGGCTCTACTACTGTACTGTCAAGTGCA 240
    |||||||

QY 609 CTTTGATGAGGGGAGGAGGCTGTCTACTGAAGCTGAGCTTGCTGATGATGATGCTGGC 668
    |||||||
Db 241 CTTTGATGAGGGGAGGAGGCTGTCTACTGAAGCTGAGCTTGCTGATGATGATGCTGGC 300
    |||||||

QY 669 CTTGGCGCTGCTGAGGAATTCAGCCACTGCGGCCAGTTCCCTCGGG 717
    |||||||
Db 301 CTTGGCGCTGCTGAGGAATTCAGCCACTGCGGCCAGTTCCCTCGGG 349
    |||||||

```

Search completed: May 8, 2003, 13:50:17  
Job time : 1862 secs

**THIS PAGE BLANK (uspro)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 10:15:30 ; Search time 342 Seconds  
(without alignments)  
9040.917 Million cell updates/sec

Title: US-09-245-198a-3  
Perfect score: 1373  
Sequence: 1 atgcatgttgtagcttga.....gacaaatgtgtaaatg 1373

Scoring table: OLIGO\_MNC  
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N.GeneSeq\_101002.\*

```
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1373	100.0	1373	19	AAV18600
2	1247	90.8	1364	24	ABK34881
3	1172	85.4	1353	21	AAA49717
4	1172	85.4	1353	24	ABK40255
5	1172	85.4	1421	20	AAV56000
6	958	69.8	1236	19	AAV47613
7	958	69.8	1236	22	AAV40350
8	784	57.1	1030	20	AAV23424
9	625	45.5	898	22	AAV03964

10	179	13.0	195	24	ABK29540	Colon adenocarcino
11	76	5.5	282	16	AAV22190	Human gene signatu
12	60	4.4	60	24	ABN41049	Human spliced tran
13	60	4.4	60	24	ABN58591	Human spliced tran
14	60	4.4	60	24	ABN58593	Human spliced tran
15	60	4.4	60	24	ABN58848	Human spliced tran
16	60	4.4	60	24	ABN58849	Human spliced tran
17	50	3.6	50	20	AAV56002	Human tumour necro
18	50	3.6	50	24	AAA49733	Human pO207 DNA p
19	50	3.6	50	24	ABK40292	Oligonucleotide pr
20	46	3.4	701	20	AAV23425	Mouse TNRL3 DNA.
21	46	3.4	1168	19	AAV18599	Mus musculus tumou
22	26	1.9	26	24	ABK40356	Reverse PCR primer
23	26	1.9	147	16	AAV19717	Human gene signatu
24	26	1.9	153	24	AAV69032	Activated T-cell d
25	26	1.9	281	22	AAV24709	Human ovarian PCR-
26	26	1.9	386	24	ABQ60530	Human colon cancer
27	26	1.9	391	22	AAH83337	Human ovarian tumo
28	26	1.9	439	24	ABN96246	Gene #2744 used to
29	26	1.9	439	24	ABL62690	Colon adenocarcino
30	26	1.9	452	21	AAV06922	Human secreted pro
31	26	1.9	483	24	ABK55088	Human colon cancer
32	26	1.9	486	21	AAZ51563	Human hypoxia resp
33	26	1.9	487	24	ABN94231	Gene #729 used to
34	26	1.9	516	21	AAV01272	Human secreted pro
35	26	1.9	531	24	ABK55407	Human colon cancer
36	26	1.9	540	21	AAV01271	Human secreted pro
37	26	1.9	570	21	AAV7898	Human cancer assoc
38	26	1.9	580	21	AAZ51562	Human hypoxia resp
39	25	1.8	400	24	ABQ58318	Human colon cancer
40	25	1.8	626	23	AAV64426	DNA encoding novel
41	24	1.7	24	20	AAV23450	Human TNRL3 RACE p
42	24	1.7	40	20	AAV23451	Human TNRL3 RACE p
43	24	1.7	40	20	AAV56003	Human tumour necro
44	23	1.7	23	24	ABK40355	Probe for gene amp
45	22	1.6	38	20	AAV56004	Human tumour necro

## ALIGNMENTS

RESULT 1	
AAV18600	
ID	AAV18600 standard; cDNA; 1373 BP.
AC	AAV18600;
XX	
XX	21-JUL-1998 (first entry)
DT	
XX	
XX	Home sapiens tumour necrosis factor related ligand (TNFRL) gene.
DE	
XX	
XX	TNRL; tumour necrosis factor related ligand; tnfr; treatment;
KW	cancer; autoimmune disease; immune system; stimulation; suppression;
KW	graft rejection; ds.
XX	
OS	Home sapiens.
XX	
FT	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	1..852
FT	/*tag= a
FT	/note= "tumour necrosis factor related ligand"
PN	
XX	WO9805783-A1.
PD	
XX	12-FEB-1998.
XX	
PF	07-AUG-1997;
XX	97WO-US13945.
XX	
PR	18-MAR-1997;
PR	97US-0040820.
PR	07-AUG-1996;
PR	96US-0023541.
XX	18-OCT-1996;
XX	96US-0028515.
PA	(BIOT ) BIOGEN INC.

PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.

XX Browning JL, Chicheportliche Y;

XX WPI; 1998-145619/13.

DR P-PSDB; AAM47525.

XX Tumour necrosis factor related ligand - useful for, e.g. treating  
PT cancer, auto-immune disease and immune responses to tissue grafts

PS Claim 2: Pages 48-50; 69pp; English.

XX The sequence is that encoding human tumour necrosis factor related  
CC ligand (TREML). TREML or active fragments can be included with a  
CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
CC diseases or immune responses to tissue grafts, or to stimulate or  
CC suppress the immune system. It is useful to screen for TREML  
CC receptors, by labelling with a detectable label and screening  
CC compositions for binding. Agents interfering with TREML-receptor  
CC binding can also be screened for, can then be administered,  
CC optionally with interferon gamma, to induce cell death or  
CC treat, suppress or alter immune responses (especially involving human  
CC adenocarcinoma cells) involving a signal pathway between TREML and its  
CC receptor. The DNA sequence can be used in gene therapy for  
CC TREML-related disorders in mammals (especially humans), e.g. tumours,  
CC autoimmune and inflammatory diseases or inherited genetic disorders,  
CC by introducing into cells, and expressing, therapeutically effective  
CC amounts of a vector, e.g. a virus comprising a gene encoding TREML.  
CC It may also be of use in the preparation of probe probes for  
CC screening natural/synthetic DNAs for TREML-encoding sequences  
CC and for antisense therapy.

SQ Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other:

Query Match 100.0%; Score 1373; DB 19; Length 1373;  
Best Local Similarity 100.0%; Pred. No. 0;

Matches 1373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCATTTGTAGACTTTGAAATTTCCGCGCGGCTCCCTCCCGCATCTCTGCGG 60  
DB 1 ATGCATTTGTAGACTTTGAAATTTCCGCGCGGCTCCCTCCCGCATCTCTGCGG 60  
QY 61 TCCGCGGATGGGGGGGGGCTGAGGACAGCCCGCCCGCCATGGCGCGCGTGG 120  
DB 61 TCCGCGGATGGGGGGGGGCTGAGGACAGCCCGCCCGCCATGGCGCGCGTGG 120  
QY 121 AGCCAGAGGCGGAGGGGGCGCGGGGGAGCCGGGACCGCTGCTGCTCCGCTCGC 180  
DB 121 AGCCAGAGGCGGAGGGGGCGCGGGGGAGCCGGGACCGCTGCTGCTCCGCTCGC 180  
QY 181 CTGGGCGGCGGCGCTGGGCGCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
DB 181 CTGGGCGGCGGCGCTGGGCGCTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
QY 241 AGCGGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
DB 241 AGCGGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300  
QY 301 CAGAGCCGCTGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
DB 301 CAGAGCCGCTGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 361 AACCGACTAGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
DB 361 AACCGACTAGTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420  
QY 421 GCGATCCGACGCCCATTTATGAAGTTATCATCAAGCTGGAAGGAGGAGGAGGAGGAG 480  
DB 421 GCGATCCGACGCCCATTTATGAAGTTATCATCAAGCTGGAAGGAGGAGGAGGAGGAG 480  
QY 481 GTGAGCGGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 540  
DB 481 GTGAGCGGAGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 540

QY 541 TACAAACCGCAGATCGGGAGTTTATAGTACACCGGGGCTGCTACTACTACTACTACT 600  
DB 541 TACAAACCGCAGATCGGGAGTTTATAGTACACCGGGGCTGCTACTACTACTACTACT 600  
QY 601 CAGGTGCACTTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
DB 601 CAGGTGCACTTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660  
QY 661 GTGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
DB 661 GTGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
QY 721 CAGCTCCGCTTGGCAGAGTGTGAGGCTGTTGGGCTTGGGCTTGGGCTTGGGCTTGGG 780  
DB 721 CAGCTCCGCTTGGCAGAGTGTGAGGCTGTTGGGCTTGGGCTTGGGCTTGGGCTTGGG 780  
QY 781 ATCCGACACCTTCCCTGGGCGCCATCTCAAGGCTGCGCCCTTCTCACTACTTGGAGTC 840  
DB 781 ATCCGACACCTTCCCTGGGCGCCATCTCAAGGCTGCGCCCTTCTCACTACTTGGAGTC 840  
QY 841 TTCCAGTTCACTAGAGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
DB 841 TTCCAGTTCACTAGAGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900  
QY 901 CGACAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
DB 901 CGACAGCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960  
QY 961 CCGTGGCGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
DB 961 CCGTGGCGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020  
QY 1021 ACAGTATCCCACTTTATTTATTAACACTCCCGCCAGCTGCTGCTGCTGCTGCTGCTG 1080  
DB 1021 ACAGTATCCCACTTTATTTATTAACACTCCCGCCAGCTGCTGCTGCTGCTGCTGCTG 1080  
QY 1081 CCCAATCCCTGACCTTTGAGGCGCCAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
DB 1081 CCCAATCCCTGACCTTTGAGGCGCCAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140  
QY 1141 CAGGCAATTTGTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
DB 1141 CAGGCAATTTGTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
QY 1201 GCACATAAGAGGGGCTGACCTGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260  
DB 1201 GCACATAAGAGGGGCTGACCTGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260  
QY 1261 TTCCCAATTTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
DB 1261 TTCCCAATTTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
QY 1321 TTTTAAACAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1373  
DB 1321 TTTTAAACAGATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1373

RESULT 2  
ABK34881  
ID ABK34881 standard; cDNA; 1364 BP.  
XX  
XX ABK34881;  
AC  
AC  
DT 08-MAY-2002 (first entry)  
XX  
DE Human cDNA encoding secreted protein #19.  
XX  
XX Human; secreted protein; gene; ss; nutritional supplement; haemophilia;  
KW viral infection; bacterial infection; fungal infection; diabetes; asthma;  
KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;  
KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;  
KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;





AA	AAA49717	
ID	AAA49717	standard; cDNA; 1353 BP.
XX		
AC	AAA49717;	
XX		
DT	25-SEP-2000	(first entry)
XX		
DE	Human PRO207 cDNA clone DNA30879-1152.	
XX		
KW	PRO207; human; antitumour; tumour; therapy; cytostatic;	
KW	breast cancer; ovarian cancer; renal cancer; colorectal cancer;	
KW	uterine cancer; prostate cancer; lung cancer; bladder cancer;	
KW	central nervous system cancer; melanoma; leukaemia; neoplasm; ss.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	58..807
FT		/*tag- a
FT	sig_peptide	58..177
FT		/*tag- b
FT	mat_peptide	178..804
FT		/*tag- c
XX		
PN	MO200037638-A2.	
PD		
XX	29-JUN-2000.	
XX		
FE	02-DEC-1999;	99MO-US28565.
XX		
PR	22-DEC-1998;	98US-0113296.
PR	08-MAR-1999;	99MO-US05028.
PR	21-APR-1999;	99US-0130232.
PR	28-APR-1999;	98US-0131445.
PR	14-MAY-1999;	99US-0134287.
PR	20-JUL-1999;	99US-0144758.
PR	26-JUL-1999;	99US-0145698.
PR	15-SEP-1999;	99MO-US21090.
PR	15-SEP-1999;	99MO-US21447.
XX		
PA	(GETH ) GENENTECH INC.	
PI	Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;	
PI	Napier MA, Pittl RM, Wood WI;	
XX		
DR	WPI: 2000-442668/38.	
DR	P-PSDB: AAY95338.	
XX		
PT		
PT	PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or	
PT	PRO866	
XX		
PS	Claim 20; Fig 3; 172pp; English.	
XX		
CC	The present sequence is that of cDNA clone DNA30879-1152	
CC	(ATCC 209358) encoding human PRO207 (see AAY95338), which shows	
CC	homology to several members of the tumour necrosis factor family,	
CC	especially human lymphotoxin (23.4%). The cDNA was identified in a	
CC	fetal kidney cDNA library following identification of an expressed	
CC	sequence tag with homology to human Apo-2 ligand. A claimed method	
CC	for inhibiting the growth of a tumour cell comprises exposing the	
CC	tumour cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224,	
CC	PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866 (see	
CC	AAY95337-49), their agonists or chimeric polypeptides incorporating	
CC	them. The tumour is especially a cancer selected from breast,	
CC	ovarian, renal, colorectal, uterine, prostate, lung, bladder and	
CC	central nervous system cancer, melanoma and leukaemia. Nucleic	
CC	acids encoding PRO179 etc. are used in the recombinant production	
CC	of the antitumour polypeptides.	
XX		
Q0	Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;	

Query Match		85.4%	Score 1172	DB 21	Length 1353
Best Locality		Similarity 99.8%	Pred. No 0		
Matches 1322		Conservative 0	Mismatches 3	Indels 0	Gaps
QY	CGATCCCTCGGGTCCCGGGATGGGGGGCGGTGAGCGACAGCGCCCGCCCATG	108			
Db	1CGATCCCTCGGGTCCCGGGATGGGGGGCGGTGAGCGACAGCGCCCGCCCATG	60			
QY	GGCGGCGGTGCGAGCGAGGCGGAGGGGGGCGCGGGGGAGCGCGGGCAACCGCTGCTG	168			
Db	61GGCGGCGGTGCGAGCGAGGCGGAGGGGGGCGCGGGGGAGCGCGGGCAACCGCTGCTG	120			
QY	GTCCCGCTCGGCTGGGCTGGGCTGGCGCTGGCTCGGCTCGGCTCTGCTGGCGGTG	228			
Db	121GTCCCGCTCGGCTGGGCTGGGCTGGCGCTGGCGCTGGCTCGGCTCTGCTGGCGGTG	180			
QY	GTCAGTTTGGGGAGCCGGGCACTGCCTTCCGCCAGAGAGCTTCCAGAGAGAGTGGTG	288			
Db	181GTCAGTTTGGGGAGCCGGGCACTGCCTTCCGCCAGAGAGCTTCCAGAGAGAGTGGTG	240			
QY	GCAGAGAGAGACACAGACCCGTCGGAACTGATCCCGACAGACAGAAAGCCAGATCTT	348			
Db	241GCAGAGAGAGACACAGACCCGTCGGAACTGATCCCGACAGACAGAAAGCCAGATCTT	300			
QY	GGCGCTTTCGTAACCGACTAGTTGCGCTCGCAGAAAGTCACTAAAGCGGAAACA	408			
Db	301GGCGCTTTCGTAACCGACTAGTTGCGCTCGCAGAAAGTCACTAAAGCGGAAACA	360			
QY	CGGGCTCGAAGAGGATCGAGCCCATTAATGAAGTTATCATCAGACCTGGACAGAGCA	468			
Db	361CGGGCTCGAAGAGGATCGAGCCCATTAATGAAGTTATCATCAGACCTGGACAGAGCA	420			
QY	GGCGAGGAGGTGTGAGAGGGAGACAGTACGTAGGCTGGAGAGAAAGCCAGATCAACAGCTCC	528			
Db	421GGCGAGGAGGTGTGAGAGGGAGACAGTACGTAGGCTGGAGAGAAAGCCAGATCAACAGCTCC	480			
QY	AGCCCTTCGCGCTACAAACCGCGCAATGGGGAGTTTATAGTCACCCGGGCTGGGCTTAC	588			
Db	481AGCCCTTCGCGCTACAAACCGCGCAATGGGGAGTTTATAGTCACCCGGGCTGGGCTTAC	540			
QY	TACCTGTACTCTCAGGTGCACTTTTGATGAGGGGAAAGGCTGTACCTGAAGCTGCACTTG	648			
Db	541TACCTGTACTCTCAGGTGCACTTTTGATGAGGGGAAAGGCTGTACCTGAAGCTGCACTTG	600			
QY	CTGTGGATGGTGTGTGTCGGCCCTGCGCTGAGAGAAATTCACCACTCGCGGCACT	708			
Db	601CTGTGGATGGTGTGTGTCGGCCCTGCGCTGAGAGAAATTCACCACTCGCGGCACT	660			
QY	TCCCTCGGGGCCCCAGCTCCGCGCTGTGCAAGTGTCTGGGCTGTGGCCCTCGCGCAGAG	768			
Db	661TCCCTCGGGGCCCCAGCTCCGCGCTGTGCAAGTGTCTGGGCTGTGGCCCTCGCGCAGAG	720			
QY	TCCTCCCTGGGATCGGACACCCTCCCTGGGGCCATCTCAAGGCTGCCCCCTTCTCAAC	828			
Db	721TCCTCCCTGGGATCGGATCCGACACCCTCCCTGGGGCCATCTCAAGGCTGCCCCCTTCTCAAC	780			
QY	TACTTGGAGACTTCCAGAGTTCACTAGAGGGGCCCCGTGTCCCGCAGTCGCCAGGCT	888			
Db	781TACTTGGAGACTTCCAGAGTTCACTAGAGGGGCCCCGTGTCTCCCGCAGTCGCCAGGCT	840			
QY	GGCGGCTCCCTCGACAGCTCTGGGACACCCGCTGCCCTTGCCCCACCTTCACAGCGCT	948			
Db	841GGCGGCTCCCTCGACAGCTCTGGGACACCCGCTGCCCTTGCCCCACCTTCACAGCGCT	900			
QY	CTTTGCTTCAGACTCGCCCTCCCTCTAAGAGGTGCTGTGGGCTGTTCACGTGTTTCA	1008			
Db	901CTTTGCTTCAGACTCGCCCTCCCTCTAAGAGGTGCTGTGGGCTGTTCACGTGTTTCA	960			
QY	TCCACATAAATACAGATATCCACACTCTATCTTAACTCCGCCACCGCACACTCTCA	1068			
Db	961TCCACATAAATACAGATATCCACACTCTATCTTAACTCCGCCACCGCACACTCTCA	1020			
QY	CTTCAGTACTCTCCCAATCCCTGACCCCTTTGAGAGGCCCAAGTATCTGCACTCCCCCTG	1128			



```

Db 721 TCTCCCTGGGATCCGACCCCTCCCTGGGCCCATCTCAAGCTGCCCCCTTCACACC 780
Oy 829 TACTGGAGACTTTCAGAGTTCACTGAGGGGCCCTGGTCCCGCACAGTGTCCAGGGCT 888
Db 781 TACTTGGAGACTTTCAGAGTTCACTGAGGGGCCCTGGTCCCGCACAGTGTCCAGGGCT 840
Oy 889 GCGGGCTCCCTCCGAGAGCTCTGTGGGACACCCGGTCCCTCTGACCCACCTCAGCGCT 948
Db 841 GCGGGCTCCCTCCGAGAGCTCTGTGGGACACCCGGTCCCTCTGACCCACCTCAGCGCT 900
Oy 949 CTCTTGTCCAGAGCTGCCCCCTCCCTAGAGGCTGCTGGGCTGTTCACGTGTTTCCA 1008
Db 901 CTCTTGTCCAGAGCTGCCCCCTCCCTAGAGGCTGCTGGGCTGTTCACGTGTTTCCA 960
Oy 1009 TCCCATATAATACAGATTCGACCTTATCTTACAACTCCCGCCACCGCCACTCTCCA 1068
Db 961 TCCCATATAATACAGATTCGACCTTATCTTACAACTCCCGCCACCGCCACTCTCCA 1020
Oy 1069 CCTCACTAGCTCCCAATCCCTGACCCCTTGAAGGCCCCCAAGATCTGACTCCCTCTG 1128
Db 1021 CCTCACTAGCTCCCAATCCCTGACCCCTTGAAGGCCCCCAAGATCTGACTCCCTCTG 1080
Oy 1129 GCCACAGACCCCGAGGCTGTTGTTCACTGTACTCTGTGGGCAAGAGTGGTCCAGAAG 1188
Db 1081 GCCACAGACCCCGAGGCTGTTGTTCACTGTACTCTGTGGGCAAGAGTGGTCCAGAAG 1140
Oy 1189 ACCCACTTCAGGACCTAAGAGGGGCTGACCTGCGGCGAGAAAGCAAGACTGCGG 1248
Db 1141 ACCCACTTCAGGACCTAAGAGGGGCTGACCTGCGGCGAGAAAGCAAGACTGCGG 1200
Oy 1249 CTAGGCGAGAGTTCCTCAAAATGTGAGGGCGAGAAACAGACAGCTCTCCCTTGAGAA 1308
Db 1201 CTAGGCGAGAGTTCCTCAAAATGTGAGGGCGAGAAACAGACAGCTCTCCCTTGAGAA 1260
Oy 1309 TTCCCTGTGATTTTAAACAGATATTTATTTATATGTCGCAAAATGTTGATA 1368
Db 1261 TTCCCTGTGATTTTAAACAGATATTTATTTATATGTCGCAAAATGTTGATA 1320
Oy 1369 AATGG 1373
Db 1321 AATGG 1325

RESULT 5
AAK56000
ID AAK56000 standard; DNA; 1421 BP.
XX
AC AAK56000;
XX
DT 15-JUL-1999 (first entry)
XX
DE Human tumour necrosis factor Apo-3 ligand polynucleotide sequence.
XX
KW Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;
KW NF-kappaB-dependent transcription; JNK/SAPK-dependent response;
KW cancer; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 92..841
FT /tag= a
FT /product= "Apo-3 ligand"
XX
XX MO9919490-A1.
XX
XX 22-APR-1999.
XX
XX 09-OCT-1998; 98WO-US21407.
XX
XX 17-DEC-1997; 97US-0069862.
XX
XX 10-OCT-1997; 97US-0062037.

```

```

XX (GETH ) GENENTECH INC.
XX PA
XX PI Ashkenazi AJ, Marsters SA, Pittl R;
XX DR WPI; 1999-287982/24.
XX DR P-PSDB; AAY09369.
XX
XX New human Apo-3 ligand (a tumour necrosis factor) homologue
XX
XX PS Claim 18; Fig 1; 74pp: English.
XX
XX CC The present sequence encodes a human tumour necrosis factor (TNF) and
XX CC lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
XX CC cytostatic activity. Apo-3 ligand can be used to induce apoptosis in
XX CC mammalian cancer cells, to induce NF-kappaB-dependent transcription and
XX CC to induce JNK/SAPK-dependent responses in mammalian cells.
XX
XX SQ Sequence 1421 BP; 281 A; 464 C; 404 G; 272 T; 0 other;
XX
XX Query Match 85.4%; Score 1172; DB 20; Length 1421;
XX Best Local Similarity 99.8%; Pred No. 0;
XX Matches 1322; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
Oy 49 CGATCCCTCGGGTCCCGGGATGGGGGGGGGCTGAGGACAGCAGCCCGCCCATG 108
Db 35 CGATCCCTCGGGTCCCGGGATGGGGGGGGGCTGAGGACAGCAGCCCGCCCATG 94
Oy 109 GCGGCGCGTCGAGCCAGAGGCGGAGGGGGCGCGGGGGAGCCGGGCAACCGCCGCTG 168
Db 95 GCGGCGCGTCGAGCCAGAGGCGGAGGGGGCGCGGGGGAGCCGGGCAACCGCCGCTG 154
Oy 169 GTCCCGCTCGCGCTGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 228
Db 155 GTCCCGCTCGCGCTGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 214
Oy 229 GTGAGTTGGGGAGCCGGGATCGCTGTCGCCAGAGAGCTGCGCCAGAGAGCTGTG 288
Db 215 GTGAGTTGGGGAGCCGGGATCGCTGTCGCCAGAGAGCTGCGCCAGAGAGCTGTG 274
Oy 289 GCAGAGAGAGGACGAGACCGCTGGGAATGATCCGACAGCAAGAAAGCCAGATCC 348
Db 275 GCAGAGAGAGGACGAGACCGCTGGGAATGATCCGACAGCAAGAAAGCCAGATCC 334
Oy 349 GCGCCTTCTGAAACGAGTATGTCGCGCTCGCAGAAAGTGCACCTAAAGCGCGAAACA 408
Db 335 GCGCCTTCTGAAACGAGTATGTCGCGCTCGCAGAAAGTGCACCTAAAGCGCGAAACA 394
Oy 409 CGGCTCGAAGAGCGATCGCAGCCATTTAATGAATTCATCCAGACCTGGACAGCGA 468
Db 395 CGGCTCGAAGAGCGATCGCAGCCATTTAATGAATTCATCCAGACCTGGACAGCGA 454
Oy 469 GCGCAGGACAGTGTGACGGGACAGTGAAGTGTGGGAGGAAGCCAGATTCACACTGC 528
Db 455 GCGCAGGACAGTGTGACGGGACAGTGAAGTGTGGGAGGAAGCCAGATTCACACTGC 514
Oy 529 AGCCCTCTGCGCTACCAACCGCCAGATCGGGGAGTTTATAGTACCCCGGCTGGCTCTAC 588
Db 515 AGCCCTCTGCGCTACCAACCGCCAGATCGGGGAGTTTATAGTACCCCGGCTGGCTCTAC 574
Oy 589 TACTGTACTGTGAGTGCATTTGATGAGGGAGAGCTGTCTACTGAAGCTGCACTTG 648
Db 575 TACTGTACTGTGAGTGCATTTGATGAGGGAGAGCTGTCTACTGAAGCTGCACTTG 634
Oy 649 CTGCTGATGATGTGCTGGGCGCTGCGCTCGGAGGAATTCACCACTGGGCGCATG 708
Db 635 CTGCTGATGATGTGCTGGGCGCTGCGCTCGGAGGAATTCACCACTGGGCGCATG 694
Oy 709 TCCCTGCGGCGCCAGCTCGGCTCTGCGAGGTGTGCGGCTGTTGGCCCTGCGGCGAGGG 768
Db 695 TCCCTGCGGCGCCAGCTCGGCTCTGCGAGGTGTGCGGCTGTTGGCCCTGCGGCGAGGG 754
Oy 769 TCTCTCCCTGGGATCCGACACCTCCCTCGGGCCCATCTCAAGGCTGCCCCCTTCTCAC 828

```

Db	755	TCCTCCCTGGGAGATCCGCAACCCCTCCCTGGGGCCAAATCTCAAGAGGTGGCCCTTCTCTCAAC	814
Qy	829	TACTTTCGGACTCTTCCAGGTTCACTAGAGGGCCCTGGTCTCCCAACAGTGTCCAGGCT	888
Db	815	TACTTTCGGACTCTTCCAGGTTCACTAGAGGGGGCCCTGGTCTCCCAACAGTGTCCAGGCT	874
Qy	889	GCCGGCTCCCTCTGACAGCTCTCTGGGCAACCGGTCGCCCTCTGGCCCAACCTCGACGGCT	948
Db	875	GCCGGCTCCCTCTGACAGCTCTCTGGGCAACCGGTCGCCCTCTGGCCCAACCTCGACGGCT	934
Qy	949	CTTTGTCTCCAGACCTGCCCCCTCCCTCTAGAGGCTGCTGGGCGCTTTCAGCTGTTTTCCA	1008
Db	935	CTTTGTCTCCAGACCTGCCCCCTCCCTCTAGAGGCTGCTGGGCGCTTTCAGCTGTTTTCCA	994
Qy	1009	TCCACATTTAAATACAGATTAATCCCACTCTTATCTTCAACTCCGCCCAACGCCCACTGCCA	1068
Db	995	TCCACATTTAAATACAGATTAATCCCACTCTTATCTTCAACTCCGCCCAACGCCCACTGCCA	1054
Qy	1069	CTCTCAGTAGTCCCCCAATCCCTCAGACCCCTTTGAGGGCCCCAGTGAATCTGCACTGCCCTG	1128
Db	1055	CTCTCAGTAGTCCCCCAATCCCTCAGACCCCTTTGAGGGCCCCAGTGAATCTGCACTGCCCTG	1114
Qy	1129	GCCACAGACCCCCAGGGCAATTTGGTCACTGTAACACTGTCGTGGGCCAAGATATGGTCCAAAG	1188
Db	1115	GCCACAGACCCCCAGGGCAATTTGGTCACTGTAACACTGTCGTGGGCCAAGATATGGTCCAAAG	1174
Qy	1189	ACCCCACTTCAGGCACTAAGAGAGGGGCTGGACCTGCGCGAGAGAACCCAAAGACTGGGC	1248
Db	1175	ACCCCACTTCAGGCACTAAGAGAGGGGCTGGACCTGCGCGAGAGAACCCAAAGACTGGGC	1234
Qy	1249	CTAGGCCAGGAGTTCCCAATTTGAGAGGGCGGAGAAACAAGACAGCTCTCTCTTGAGAA	1308
Db	1235	CTAGGCCAGGAGTTCCCAATTTGAGAGGGCGGAGAAACAAGACAGCTCTCTCTTGAGAA	1294
Qy	1309	TTCCCTCTGGATTTTTTAAACAAGATATTAATTTTATATATTTGACAAATGTGATA	1368
Db	1295	TTCCCTCTGGATTTTTTAAACAAGATATTAATTTTATATATTTGACAAATGTGATA	1354
Qy	1369	AATGG 1373	
Db	1355	AATGG 1359	

RESULT 6			
AAV47613			
ID	AAV47613	standard; cDNA; 1236 BP.	
XX			
AC	AAV47613;		
XX			
DE	27-OCT-1998	(first entry)	
XX			
DE	TNF related endothelium proliferative agent gene.		
XX			
KW	ss: TNF; endothelium proliferative agent; TREPA: wound healing; cancer		
XX	tissue grafting; vascularisation; apoptosis; autoimmune; birth control		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..750	
FT		/*tag= a	
FT		/product= "TREPA"	
XX			
PN	W09835061-A2.		
XX			
PD	13-AUG-1998.		
XX			
PF	12-FEB-1998;	98WO-US02859.	
XX			
PR	10-FEB-1998;	98US-0021706.	
PR	12-FEB-1997;	97US-0798692.	
XX			

PA (ABBO ) ABBOTT LAB.  
XX  
XX  
PI Wiley SR;  
XX  
XX WPI: 1998-447255/38.  
DR  
DR P-ESDB; AAM29745.  
DR  
XX  
XX  
PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
PT treatment of autoimmune disease, tumours and inflammation  
XX  
XX  
PS Claim 11; Page 123-4; 142pp: English.

CC The TRP-related endothellium proliferative agent (TRPP), or its  
CC activators or agonists, are used to treat a deficit of TRPP, e.g. to  
CC promote wound healing or tissue grafting, by promoting vasculatisation,  
CC also to induce apoptosis for treating cancer and eliminating autoreactive  
CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
CC TRPP peptides can also be used to target cytotoxic agents or for  
CC affinity isolation of the corresponding receptor, the nucleic acid for  
CC which can be used to transform tumour cells to render them more  
CC responsive to TRPP and to screen for TRPP mimics.  
CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
CC TRPP-associated diseases, e.g. tumours and metastases (by inhibiting  
CC vasculatisation), inflammation or a wide range of autoimmune conditions,  
CC conditions involving abnormal stimulation of epithelial cells (e.g.  
CC atherosclerosis), for birth control (inhibiting ovulation and placental  
CC formation) or other angiogenic conditions (e.g. ulcers).  
XX

Query Match	69.8%	Score 958	DB 19	Length 1236
Best Local Similarity	99.6%	Pred. NO. 0		
Matches 1208	Conservative	0	Mismatches 5	Indels 0
				Gaps 0

OY	129	GCAGAGGGGGCGCCGGGGGAGCCGGGACCGCCCTCTGTGTCCCGCTCGCGCTGGGCT	188
Db	24	GCGGAGGGGGGCGCGGGGGGAGCGGGGACCGCCCTCTGTGTCCCGCTCGCGCTGGGCT	83
OY	189	GGGCGCTGGCGCTGGCGCTCTCTCGGCGCTCTCTGTGGCGTGTCAAGTTTGGGAGCCGGG	248
Db	84	GGGCGTGGCGCTGGCGCTCTCTCGGCGCTCTCTGTGGCGTGTCAAGTTTGGGAGCCGGG	143
OY	249	ATCGCTGTCCGCCAGAGACCTGTGCCAGAGAGACTGGTGGCAGAGAGAGACCGAGACC	308
Db	144	ATCGCTGTCCGCCAGAGAGCTGTGCCAGAGAGACTGGTGGCAGAGAGAGACCGAGACC	203
OY	309	GTCGGAACTGAATCCCGACAGACAGAAAGCCAGATCTCTGGCCTTTCTGAACGACT	368
Db	204	GTCGGAACTGAATCCCGACAGAGAAAGCCAGATCTCTGGCCTTTCTGAACGACT	263
OY	369	AGTTGGGCTCGCGAAGTGTACCTTAAAGCGCGGAGAAACCGGGCTCGAAGAGGATCGC	428
Db	264	AGTTGGGCTCGAAGAGATGTACCTTAAAGCGCGGAGAAACCGGGCTCGAAGAGGATCGC	323
OY	429	AGCCATTATGAAGTTCATCCACGACCTGGACAGAGAGGAGCGAGCGAGGTGGACG	488
Db	324	AGCCATTATGAAGTTCATCCACGACCTGGACAGAGAGGAGCGAGCGAGGTGGACG	383
OY	489	GACAGTAGTGGCTGGGAGGAGCCAGATTAACAGCTCCAGCCCTCTGCGCTACAACG	548
Db	384	GACAGTAGTGGCTGGGAGGAGCCAGATTAACAGCTCCAGCCCTCTGCGCTACAACG	443
OY	549	CCAGATCGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACTGTACTGTCAAGTGA	608
Db	444	CCAGATCGGGAGTTTATAGTCACCCGGGCTGGGCTCTACTACTGTACTGTCAAGTGA	503
OY	609	CTTTGATAGGGGAGAGGCTGTCTACTCGAAGCTGGACTGTGTGGATGGTGGCTGGC	668
Db	504	CTTTGATAGGGGAGAGGCTGTCTACTCGAAGCTGGACTGTGTGGATGGTGGCTGGC	553
OY	669	CTGTGGCTCGCTGGAGGAATTCACGCACCTGCGGCAAGTTCCCTCGGGGCCACAGCTCG	728
Db	564	CTGTGGCTCGCTGGAGGAATTCACGCACCTGCGGCAAGTTCCCTCGGGGCCACAGCTCG	623

QY	729	CCTGTGACAGTGTCTCAGGGTGTTGGCCCTGTGGGCACAGAGTCCTCCCGGGAATCCGCAC	788
Db	624	CTCTGTGCAGAGTCTCTGGGCTGTTGGCCCTGTGGGCACAGAGTCCTCCCGGGAATCCGCAC	683
QY	789	CTCTGCCCTGGGCCCATTCTCAAGGCTGCCCCCTTCTCTCACTACTTGGACTCTTCCAGGT	848
Db	684	CTCTGCCCTGGGCCCATTCTCAAGGCTGCCCCCTTCTCTCACTACTTGGACTCTTCCAGGT	743
QY	849	TCACTAGAGGGGCCCTGTGTCTCCGCACAGTCGTCCAGAGCTGCGGCTCCCTCGACAGCT	908
Db	744	TCACTAGAGGGGCCCTGTGTCTCCGCACAGTCGTCCAGAGCTGCGGCTCCCTCGACAGCT	803
QY	909	CTCTGGGACACCGGTCCTCCCTGTGCCCCCACCCCTCAAGCCGCTGTTGGCTCAGAGCTGACCC	968
Db	804	CTCTGGGACACCGGTCCTCCCTGTGCCCCCACCCCTCAAGCCGCTGTTGGCTCAGAGCTGACCC	863
QY	969	TCCCTTAGAAGGCTGCTGGGCTGTCTCAGCTGTTTTCCATCCCATMAATACAGTAAT	1028
Db	864	TCCCTTAGAAGGCTGCTGGGCTGTCTCAGCTGTTTTCCATCCCATMAATACAGTAAT	923
QY	1029	CCCACTCTTACTTTAAACTCCGCCACCGGCCACTCTCACCCTCACTAGCTCCCCATTC	1088
Db	924	CCCACTCTTACTTTAAACAACACCCACGCGCCACTCTCACCCTCACTAGCTCCCCATTC	983
QY	1089	CTGACCCCTTTAGAGGCCGCCAGTAGATCTCGACTCCGCCCTGGGCCACAGACCCCGAGGCAT	1148
Db	984	CTGACCCCTTTAGAGGCCGCCAGTAGATCTCGACTCCGCCCTGGGCCACAGACCCCGAGGCAT	1043
QY	1149	TGTGTTTACTGTACTCTGTGTGGGCAGAAGATGGGTCCAGAAAGACCCCACTTCAAGCATAAG	1208
Db	1044	TGTGTTTACTGTACTCTGTGTGGGCAGAAGATGGGTCCAGAAAGACCCCACTTCAAGCATAAG	1103
QY	1209	AGGGGCTGAGACCTGGGGGCGAGAAAGCAAAGAGACTGGGGCTTAGGGCCAGAGTTCCCAA	1268
Db	1104	AGGGGCTGAGACCTGGGGGCGAGAAAGCAAAGAGACTGGGGCTTAGGGCCAGAGTTCCCAA	1163
QY	1269	TGTGAGGGGCGAAMAACAAGACAAGCTCCTCCCTTGAGAAATTCCTGTGATTTTTTAAAA	1328
Db	1164	TGTGAGGGGCGAAMAACAAGACAAGCTCCTCCCTTGAGAAATTCCTGTGATTTTTTAAAA	1223
QY	1329	CAGATATTATTATTT 1341	
Db	1224	CAGATATTATTATTT 1236	
 RESULT 7 AAD04350 standard; cDNA: 1236 BP.			
ID	AAD04350	standard; cDNA: 1236 BP.	
XX	XX		
AC	AAD04350;		
XX	XX		
DT	04-JUL-2001	(first entry)	
XX	XX		
DE	Human TREPA (TNF related endothelium proliferative agent) cDNA.		
XX	XX		
KW	Human; tumour necrosis factor; TNF; angiogenesis; wound healing; TREPA; TNF related endothelium proliferative agent; tumour; metastasis		
KX	grafting; vulnerability; ss.		
OS	Homo sapiens.		
FH	XX		
FT	Key	Location/Qualifiers	
FT	CDS	1..750	
FT	FT	/tag= a	
FT	XX	/product= "Human TREPA (TNF related endothelium proliferative agent)".	
PN	US6207642-BI.		
PD	27-MAR-2001.		
PF	26-JUN-1998; 980S-0105343.		

XX	12-FEB-1997;	97US-0798692.
PR	10-FEB-1998;	98US-0021706.
XX	(ABBO ) ABBOTT LAB.	
XX	Wiley SR:	
PI	WPI: 2001-280760/29.	
DR	P-PSDB: AAE00891.	
XX	Inducing angiogenesis in mammal at desired sites for promoting wound	
PT	healing, by administering soluble fragment of extracellular domain of	
PT	tumor necrosis factor related endothelium proliferative agent protein	
PT	-	
XX	Example 2; Column 73-74; 53pp; English.	
PS	The present invention relates to extracellular signal molecules,	
CC	particularly members of tumour necrosis factor (TNF) family molecules	
CC	designated as TREPA (TNF related endothelium proliferative agent).	
CC	Soluble biologically active TREPA are used to treat TREPA-associated	
CC	diseases, tumours or metastases. TREPA is used for inducing angiogenesis	
CC	in human for promoting wound healing and for vascularising grafted	
CC	tissue for successful grafting and to promote tissue grafts.	
CC	The present sequence is a cDNA clone ID #690050 encoding human TREPA.	
XX	Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other:	
SO		
	Query Match	69.8%; Score 958; DB 22; Length 1236;
	Best Local Similarity	99.6%; Pred. No. 0;
	Matches 1208; Conservative	0; Mismatches 5; Indels 0; Gaps 0.
OY	129 GCGGAGGGGGGCGGGGGGAGCCGGGACCGGCCCTGCGTGGCCGCTGGGCGCT	188
DB	24 GGGGAGGGGGGCGGGGGGAGCCGGGACCGGCCCTGCGTGGCCGCTGGGCGCT	83
OY	189 GGGCGCTGGGCGCTGGCGCTGCGCTGCGCTGCGTGGCGCGTGCAGTTTGGGAGCCGGCG	248
DB	84 GGGCGCTGGGCGCTGGCGCTGCGCTGCGCTGCGTGGCGCGTGCAGTTTGGGAGCCGGCG	143
OY	249 ATCGCTGTCGCCCGCAGAGACCTGCGCCAGAGAGAGCTGGTGGCAGAGAGAGACCCAGACCC	308
DB	144 ATCGCTGTCGCCCGCAGAGACCTGCGCCAGAGAGAGCTGGTGGCAGAGAGACCCAGACCC	203
OY	309 GTCGGAACGTGAATCCCGCAGAGAGAGAGAGAGAGATCCCTGGCGCTTCTGTAACGACT	368
DB	204 GTCGGAACGTGAATCCCGCAGAGAGAGAGAGAGAGATCCCTGGCGCTTCTGTAACGACT	263
OY	369 AGTTGGGCGCTCGCAGAGAGTGCACCTAAAGGCGGAGAAACACGCGGCTCGAAGAGCGATCGC	428
DB	264 AGTTGGGCGCTCGAAGAGAGTGCACCTAAAGGCGGAGAAACACGCGGCTCGAAGAGCGATCGC	323
OY	429 AGCCCATTTATGAAGTTTCATCCAGACCTTGAGACAGAGAGCGCAGAGAGGTGTGACGG	488
DB	324 AGCCCATTTATGAAGTTTCATCCAGACCTTGAGACAGAGAGCGCAGAGAGGTGTGACGG	383
OY	489 GACAGTGAATGAGCTGGGAGAGAGCGAAGATCAACAGCTCAGACCCCTCTGCGTACAAACGG	548
DB	384 GACAGTGAATGAGCTGGGAGAGAGCGAAGATCAACAGCTCAGACCCCTCTGCGTACAAACGG	443
OY	549 CCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTTACTACTCTGTAAGTGAAGTGA	608
DB	444 CCAGATCGGGGAGTTTATAGTCACCCGGGCTGGGCTTACTACTCTGTAAGTGAAGTGA	503
OY	609 CTTTATATGAGGGAAGGCTGTACTTGAAGCTGGACTTGGTGGTGGATGGTGTGTGCTGCG	668
DB	504 CTTTATATGAGGGAAGGCTGTACTTGAAGCTGGACTTGGTGGTGGATGGTGTGTGCTGCG	563
OY	669 CCTGGGCGGCTCGGAGGAATTCAGCCACCTGGGCGAGTTTCCCTCGGGGCCAGGTCGG	728
DB	564 CCTGGGCGGCTCGGAGGAATTCAGCCACCTGGGCGAGTTTCCCTCGGGGCCAGGTCGG	623

```
QY 729 CCTCTGACAGGTGTCTGGGCTGTGGCCCTGCGCCAGGGTCTCCCTGCGATCCGAC 788
DB 624 CCTCTGACAGGTGTCTGGGCTGTGGCCCTGCGCCAGGGTCTCCCTGCGATCCGAC 683
QY 789 CCTCTGACAGGTGTCTGGGCTGTGGCCCTGCGCCAGGGTCTCCCTGCGATCCGAC 848
DB 684 CCTCTGACAGGTGTCTGGGCTGTGGCCCTGCGCCAGGGTCTCCCTGCGATCCGAC 743
QY 849 TCACGTAGAGGGGCTGTGGGCTGTGGCCCTGCGCCAGGGTCTCCCTGCGATCCGAC 908
DB 744 TCACGTAGAGGGGCTGTGGGCTGTGGCCCTGCGCCAGGGTCTCCCTGCGATCCGAC 803
QY 909 CTCTGAGGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 968
DB 804 CTCTGAGGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 863
QY 969 TCCTCTGAGAGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGT 1028
DB 864 TCCTCTGAGAGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGT 923
QY 1029 CCCACTCTATCTTACAACTCCGCCACCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1088
DB 924 CCCACTCTATCTTACAACTCCGCCACCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983
QY 1089 CTGACCTTTTGGAGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1148
DB 984 CTGACCTTTTGGAGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043
QY 1149 TGTGTTCACTGTACTGTGTGTGGGCAAGATGGTTCAGAAAGCCCACTTTCAGGCACTAG 1208
DB 1044 TGTGTTCACTGTACTGTGTGTGGGCAAGATGGTTCAGAAAGCCCACTTTCAGGCACTAG 1103
QY 1209 AGGGCTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1268
DB 1104 AGGGCTGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1163
QY 1269 TGTGAGGGGCGAGAAACAGAACAGCTCTCCCTTGTAGAAATCCCTGTGTGATTTTAA 1328
DB 1164 TGTGAGGGGCGAGAAACAGAACAGCTCTCCCTTGTAGAAATCCCTGTGTGATTTTAA 1223
QY 1329 CAGATATTATT 1341
DB 1224 CAGATATTATT 1236
RESULT 8
ID AAX23424
AAX23424 standard; DNA; 1030 BP.
AC AAX23424;
XX 18-JUN-1999 (first entry)
DE Human TNRL3 DNA.
XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
XX developmental abnormality; gestational abnormality; prostate cancer;
XX APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
XX cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
XX apoptosis; human; ss.
OS Homo sapiens.
XX Key 1..627
XX CDS /tag- a
XX /product- "TNRL3"
XX MO9911791-A2.
XX 11-MAR-1999.
XX 04-SEP-1998; 98WO-US18393.
```

```
XX 05-SEP-1997; 97US-0924634.
PR (UNIT ) UNIT WASHINGTON.
PA Chaudhary PM;
PI WPI: 1999-205191/17.
XX P-PSDB; AAM93590.
DR New Tumor Necrosis Factor family receptor polypeptides and ligands -
XX useful for diagnosis and treatment of prostate cancer and
XX developmental or gestational abnormalities
XX Example VII; Fig 13A: 156pp: English.
PS
XX This invention describes isolated Tumor Necrosis Factor (TNF) family
XX receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
XX fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
XX their active fragments. APO4 is useful for diagnosing prostate cancer
XX by determining levels of APO4 in an individual. Prostate cancer can also
XX be treated using APO4 selective binding agents linked to a therapeutic
XX moiety. APO4 polypeptides are also useful for identifying selective
XX binding agents, useful in diagnosis/treatment of disease by binding of
XX agents to the polypeptide/active fragment which is extracellular, or
XX expressed on the cell surface. The binding is preferably performed in
XX vivo. APO4 polypeptides/active fragments are also useful for screening
XX for agonists and antagonists by binding and observing the change in APO4
XX activity. Effective pharmacological agents useful in diagnosis or
XX treatment of disease are also identified using APO4 polypeptides/active
XX fragments and APO4 signal transducer molecules that specifically interact
XX with a cytoplasmic domain of APO4 and detecting a change in level of APO4
XX activity. The method is performed in vivo or in vitro. APO polypeptides
XX are all useful as immunogens for preparing antibodies. APO4 is also
XX useful for diagnosis/treatment of developmental or gestational
XX abnormalities. APO8 was transfected to human breast carcinoma cell line
XX MCF-7, and induced apoptosis.
SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;
Query Match 57.1%; Score 784; DB 20; Length 1030;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 834; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 229 GTGAGTTGGGAGCGCGGCGATCGTCCGCCAGAGCCCTGCCAGAGAGAGCTGGG 288
DB 1 GTGAGTTGGGAGCGCGGCGATCGTCCGCCAGAGCCCTGCCAGAGAGAGCTGGG 60
QY 289 GCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 348
DB 61 GCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
QY 349 GCGCTTTCTGAGACCGAGCTAGTGGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 408
DB 121 GCGCTTTCTGAGACCGAGCTAGTGGGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 409 CGGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 468
DB 181 CGGCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 469 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 528
DB 241 GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 529 AGCCCTGCGCTGACACCGCCAGATCGGAGATTATAGTACCCCGGGCTGGCTTAC 588
DB 301 AGCCCTGCGCTGACACCGCCAGATCGGAGATTATAGTACCCCGGGCTGGCTTAC 360
QY 589 TACCTGTACTGTACAGTGTACCTTTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 648
DB 361 TACCTGTACTGTACAGTGTACCTTTATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 649 CTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 708
```



```

Db 421 CTGGTGGATGGTGTGCTGGGCGCTCGCTGGAGGAAATTCACAGCCACTGCGGCCAGT 480
QY 709 TCCCTCGGGCCCGCAGCTCCGCTCTGCCAGAGTGTCTGGGGCTGTGGCCCTGCGGCCAGG 768
Db 481 TCCCTCGGGCCCGCAGCTCCGCTCTGCCAGAGTGTCTGGGGCTGTGGCCCTGCGGCCAGG 540
QY 769 TCCCTCGGGCCCGCAGCTCCGCTCTGCCAGAGTGTCTGGGGCTGTGGCCCTGCGGCCAGG 828
Db 541 TCCCTCGGGCCCGCAGCTCCGCTCTGCCAGAGTGTCTGGGGCTGTGGCCCTGCGGCCAGG 600
QY 829 TACTTGGAGCTTCTCCAGGTTCACTGAGGGGCGCTGCTCTCCCAAGTCTGCCAGCT 888
Db 601 TACTTGGAGCTTCTCCAGGTTCACTGAGGGGCGCTGCTCTCCCAAGTCTGCCAGCT 660
QY 889 GCGGCGCTCCCTCGACAGCTCTGCGGACCGGCTCCCTTGGCCCAAGTCTGCCAGCT 948
Db 661 GCGGCGCTCCCTCGACAGCTCTGCGGACCGGCTCCCTTGGCCCAAGTCTGCCAGCT 720
QY 949 CTTTGTCTCAGAGCTGCGGCTCTGAGAGGCTGCGGCTGTTTCAAGTCTGTTTCCA 1008
Db 721 CTTTGTCTCAGAGCTGCGGCTCTGAGAGGCTGCGGCTGTTTCAAGTCTGTTTCCA 780
QY 1009 TCCCACTAATATACAGTATTCCTATCTTACAACTCCCGCAGCGCCACT 1063
Db 781 TCCCACTAATATACAGTATTCCTATCTTACAACTCCCGCAGCGCCACT 835

```

RESULT 9  
 AAS03964  
 ID AAS03964 standard; DNA: 898 BP.

XX 26-SEP-2001 (first entry)

XX Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA.

XX TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;  
 XX ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 XX retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 XX rheobiosis; uveitis; macular degeneration; arthritis; rheumatism; ds;  
 XX corneal graft neovascularisation; psoriasis; metastatic condition;  
 XX malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
 XX preneoplastic condition; myocardial angiogenesis; wound granulation;  
 XX scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
 XX atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 XX peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;  
 XX fusion protein.

XX Homo sapiens.  
 XX OS  
 XX Synthetic.

XX Key Location/Qualifiers  
 XX CDS 52..873

XX FT /tag- a  
 XX FT /product- "Fusion protein comprising a growth hormone  
 XX FT leader, a leucine zipper multimerisation  
 XX FT domain, and human TWEAK extracellular  
 XX FT domain"

XX MO200145730-A2.

XX 28-JUN-2001.

XX 19-DEC-2000: 2000MO-US34755.

XX 20-DEC-1999: 99US-0172878.

XX 10-MAY-2000: 2000US-0203347.

XX (IMMUNEX CORP.

XX WILEY SR.

XX WP1: 2001-417975/44.  
 DR P-PSDB: AAU03499.  
 XX  
 PT Modulating angiogenesis in a mammal for treating diseases mediated by  
 PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  
 PT peripheral tissue, by administering antagonist or agonist of TWEAK  
 PT receptor  
 XX  
 PS Example 1: Page 39-40; 46pp; English.

XX The sequence represents a DNA from the expression vector  
 CC pDC409-LZ-TWEAK, which encodes a fusion protein comprising a growth  
 CC hormone leader, a leucine zipper multimerisation domain, and the  
 CC extracellular domain of human TWEAK. The fusion protein was used in  
 CC the isolation of human TWEAK receptor (TWEAKR)-expressing clones  
 CC from a COS cell human cDNA library. The TWEAK protein is a  
 CC member of the tumour necrosis factor (TNF) family and induces  
 CC angiogenesis. TWEAKR may therefore be used to screen for and develop  
 CC TWEAKR agonists and antagonists for the modulation of angiogenesis, to be  
 CC used in the treatment and diagnosis of human disease. The disorders  
 CC mediated by angiogenesis include ocular disorders characterised by ocular  
 CC retinoblastoma, retinopathy of prematurity, neovascular glaucoma,  
 CC rheobiosis, uveitis, macular degeneration and corneal graft  
 CC neovascularisation, and inflammatory diseases such as arthritis,  
 CC rheumatism and psoriasis. Other treatable diseases include malignant and  
 CC metastatic conditions such as sarcomas and carcinomas, benign tumours and  
 CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,  
 CC scleroderma, vascular adhesions, atherosclerotic plaque  
 CC neovascularisation, telangiectasia, wound granulation, coronary  
 CC atherosclerosis, peripheral atherosclerosis and ischaemia.  
 CC  
 XX Sequence 898 BP: 187 A; 266 C; 267 G; 178 T; 0 other:

Query Match 45.5%; Score 625; DB 22; Length 898;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-281;  
 Matches 625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 232 AGTTGGGAGACCGGGATGCTCTCCGCCAGAGAGCTGCCAGAGAGCTGTGGCA 291
Db 250 AGTTGGGAGACCGGGATGCTCTCCGCCAGAGAGCTGCCAGAGAGCTGTGGCA 309
QY 292 GAGGAGAGACGAGACCGGCTGCGAAGTCCAGACAGAAAGCCAGAGCTCGG 351
Db 310 GAGGAGAGACGAGACCGGCTGCGAAGTCCAGACAGAAAGCCAGAGCTCGG 369
QY 352 CTTTCTGAAACCACTAGTTCGGCTTCGCAAGAGTCACTTAAAGCCGGAACACGG 411
Db 370 CTTTCTGAAACCACTAGTTCGGCTTCGCAAGAGTCACTTAAAGCCGGAACACGG 429
QY 412 GCTGGAAGAGGATCGGAGCCATTATGAAGTTCACTCAGACCTGGACAGAGGAGG 471
Db 430 GCTGGAAGAGGATCGGAGCCATTATGAAGTTATCAGACCTGGACAGAGGAGG 489
QY 472 CAGCAGAGTGTGAGCGGAGACAGTGTGCTGGAGGAGACCAAGATCAACAGCTCAGC 531
Db 490 CAGCAGAGTGTGAGCGGAGACAGTGTGCTGGAGGAGACCAAGATCAACAGCTCAGC 549
QY 532 CCTGTGGCTACAAACCGCCAGATGGGGAGTTTATATCAACCGGGCTGGGCTTACTAC 591
Db 550 CCTGTGGCTACAAACCGCCAGATGGGGAGTTTATATCAACCGGGCTGGGCTTACTAC 609
QY 592 CTGTACTGTAGGTGACTTTGATGAGGGGAGGCTGTACTACGTAAGCTGAGACTTCTG 651
Db 610 CTGTACTGTAGGTGACTTTGATGAGGGGAGGCTGTACTACGTAAGCTGAGACTTCTG 669
QY 652 GTGGATGTGTGCTGTGCGCTGTGCTGAGAGAAATTCACAGCACTGGGCGCAGTTCC 711
Db 670 GTGGATGTGTGCTGTGCGCTGTGCTGAGAGAAATTCACAGCACTGGGCGCAGTTCC 729
QY 712 CTGGGCGCCAGCTCCGCTCTGCGAGGTGTGTGGCTGTGGCCCTGGGCGCAGGCTCC 771

```



DB 730 CTCGGGCCCCAGCTCCGCTCTGCCAGAGTGTCTGGGCTGTGGCCCTCGGCGGACAGGTTCC 789  
QY 772 TCCCTGGGATCCGACACCTCTCCCTGGGCCCCATCTCAAGGCTCCCTCTCTCACTAC 831  
DB 790 TCCCTGGGATCCGACACCTCTCCCTGGGCCCCATCTCAAGGCTCCCTCTCTCACTAC 849  
QY 832 TTCGACCTCTTCAGGTTCTACTGAG 856  
DB 850 TTCGACCTCTTCAGGTTCTACTGAG 874  
RESULT 10  
ABK29540  
ID ABK29540 standard; cDNA; 195 BP.  
AC ABK29540;  
DT 23-APR-2002 (first entry)  
XX  
DE Colon adenocarcinoma-specific cDNA #66.  
XX  
KM Human; colon adenocarcinoma; colon cancer; tumour; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200196389-A2.  
XX  
PD 20-DEC-2001.  
XX  
PF 07-JUN-2001; 2001WO-US18574.  
XX  
PR 09-JUN-2000; 2000US-210667P.  
XX  
PR 22-NOV-2000; 2000US-252614P.  
XX  
PA (CORI-) CORIXA CORP.  
XX  
PI Meagher MJ, King GE, Xu J, Secret H;  
XX  
DR WPI: 2002-098052/13.  
XX  
PT New isolated polynucleotide encoding a polypeptide comprising a portion  
PT of colon tumour protein, for detection, diagnosis and therapy of human  
PT colon cancer -  
XX  
PS Claim 1; Page 133; 211pp; English.  
XX  
CC The invention relates to an isolated polynucleotide (I) encoding a  
CC polypeptide (II) comprising a portion of a colon tumour protein. A new  
CC oligonucleotide (III) that hybridises to (I) is useful for  
CC determining the presence of a cancer in a patient. (II) or antigen  
CC presenting cells expressing (I) are useful for stimulating and/or  
CC expanding T cells specific for a tumour protein, by contacting T cells  
CC with (I), (II) or antigen-presenting cells that express (I), (II),  
CC or antigen presenting cells that express (II) are useful for treating  
CC a colon cancer in a patient by incubating CD4+ and/or CD8+ T cells isolated  
CC from a patient with (I), (II), or antigen presenting cells that express  
CC (II), so that T cells proliferate, and administering to the patient an  
CC effective amount of the proliferated T cells, thus inhibiting the  
CC development of a cancer in the patient. A new composition is useful for  
CC stimulating an immune response in a patient. (I) or (II) is useful in  
CC vaccines and pharmaceutical compositions for prevention and treatment of  
CC colon cancer and for the diagnosis and monitoring of the cancers. (I),  
CC (II) or an antibody against (II) is useful for detection, diagnosis and/  
CC or therapy of human colon cancer. (I) is useful as a probe or primer for  
CC nucleic acid hybridisation, and in the design and preparation of ribozyme  
CC molecules for inhibiting expression of (II) in tumour cells. ABK29475-  
CC ABK2981 represent human colon adenocarcinoma-specific cDNA sequences of  
CC the invention.  
XX  
SQ Sequence 195 BP; 49 A; 51 C; 58 G; 37 T; 0 other;

Query Match 13.0%; Score 179; DB 24; Length 195;  
Best Local Similarity 100.0%; Pred. No. 2.6e-73;

Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1146 CATTTGTTCACTGACTCTGTGGGCGACAGATGGGTCCAGAAACCCACTTCAGGCACT 1205  
DB 17 CATTTGTTCACTGACTCTGTGGGCGACAGATGGGTCCAGAAACCCACTTCAGGCACT 76  
QY 1206 AAGAGGGGCTGACCTGGGCGGACGAAAGACAGCGGCGCTAGGCGAGAGTTCCC 1265  
DB 77 AAGAGGGGCTGACCTGGGCGGACGAAAGACAGCGGCGCTAGGCGAGAGTTCCC 136  
QY 1266 AATGTGAGGGCGAGAAACAGACAGCTCCCTTGAGATTCCTCTGTGATTTT 1324  
DB 137 AATGTGAGGGCGAGAAACAGACAGCTCCCTTGAGATTCCTCTGTGATTTT 195  
RESULT 11  
AAT22190  
ID AAT22190 standard; cDNA to mRNA; 282 BP.  
AC AAT22190;  
DT 27-AUG-1996 (first entry)  
XX  
DE Human gene signature HUMGS03761.  
XX  
KM Gene signature; messenger RNA; mRNA; relative abundance; frequency;  
KM human; cloning; mapping; non-biased library; diagnosis; detection;  
KM cell typing; abnormal cell function; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO9514772-A1.  
XX  
PD 01-JUN-1995.  
XX  
PE 11-NOV-1994; 94WO-JP01916.  
XX  
PR 12-NOV-1993; 93JP-0355504.  
XX  
PA (MATS/) MATSUBARA K.  
PA (OKUB/) OKUBO K.  
XX  
PI Matsubara K, Okubo K;  
XX  
DR WPI: 1995-206931/27.  
XX  
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.  
PT for diagnosis of abnormal cell function, by preparing cDNA that  
PT reflects relative abundance of corresp. mRNA in specific human  
PT tissues  
XX  
PS Claim 1; Page 1067; 2245pp; Japanese.  
XX  
CC A single-stranded DNA (or its complementary strand or the corresp.  
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences  
CC given in AAT19001-T26837 and which is able to hybridise to part of  
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)  
CC sequences were obtained from 3'-directed cDNA libraries prepared  
CC from various human tissues; synthesis of cDNA was initiated from the  
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
CC untranslated sequence is unique to a particular mRNA species, almost  
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library  
CC is constructed so as to reflect accurately the relative abundance of  
CC different mRNAs in the particular tissue from which it was derived.  
CC The appearance frequency of a given GS in a cDNA library can be  
CC determined (esp. using primers and probes derived from the GS  
CC sequences) as a means of diagnosing abnormal cell function or for  
CC recognising different cell types.  
XX  
SQ Sequence 282 BP; 80 A; 62 C; 69 G; 66 T; 5 other;

Query Match 5.5%; Score 76; DB 16; Length 282;  
Best Local Similarity 100.0%; Pred. No. 2.8e-25;

Matches 76: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1146 CATTGTCTCACTGCTGTGGGAGGATGGCTCCAGAGACCCACTTACGACACT 1205  
|||||  
Db 36 CATTGTCTCACTGCTGTGGGAGGATGGCTCCAGAGACCCACTTACGACACT 95  
|||||

OY 1206 AAGAGGGCTGACCT 1221  
|||||

Db 96 AAGAGGGCTGACCT 111

RESULT 12  
ABN41049  
ID ABN41049 standard; DNA; 60 BP.

AC ABN41049;  
XX  
DT 15-JUL-2002 (first entry)

DE Human spliced transcript detection oligonucleotide SEQ ID NO:13797.

KW Human; mouse; rat; splice transcript; detection; RNA transcript;  
KW splice variant; transcriptome; oligonucleotide library; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200210449-A2.  
XX  
PD 07-FEB-2002.  
XX  
PE 20-JUL-2001; 2001WO-IB01903.  
XX  
PR 28-JUL-2000; 2000US-221607P.  
PR 02-MAY-2001; 2001US-287724P.  
XX  
PA (COMP-) COMPUGEN INC.  
XX  
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
XX  
DR WPI; 2002-257383/30.  
XX

PT New oligonucleotide libraries comprising oligonucleotides which  
PT selectively hybridize to mRNAs transcribed from a transcription unit of  
PT a genome, useful for detecting tissue-, pathology-, and  
PT developmental-specific genes

PS Example 1; SEQ ID 13797; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting  
CC messenger RNAs that populate a (sub-)transcriptome, where the  
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
CC transcription units that populate a genome. The library comprises  
CC several oligonucleotides, each capable of hybridizing selectively to a  
CC set of messenger RNAs transcribed from a given transcription unit of  
CC the genome, which encodes one or more messenger RNA splice variants.  
CC The oligonucleotide libraries are useful for detecting mRNAs from a  
CC biological sample, in expression profiling studies, in qualitatively or  
CC quantitatively characterizing the corresponding transcriptome, and in  
CC detecting RNA transcripts and splice variants of human or animal  
CC transcriptomes. The libraries may also be used as specialised mini  
CC libraries to detect transcripts of a sub-transcriptome under a  
CC particular biological or pathological state, and so allowing the  
CC detection of tissue- and pathology-specific genes such as those genes  
CC only expressed in specific tissue under a specific pathological  
CC condition; to detect developmental specific genes; and to detect RNA  
CC transcripts and splice variants of a transcriptome of a patient suffering  
CC from a particular disorder. ABN27253 to ABN55589 represent  
CC oligonucleotide sequences from rats, humans and mice, which are used in  
CC the exemplification of the present invention.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

SO Sequence 60 BP; 17 A; 15 C; 15 G; 13 T; 0 other;

Query Match 4.4%; Score 60; DB 24; Length 60;  
Best Local Similarity 100.0%; Pred. No. 8 6e-18;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1263 CCCAATGTGAGGGCGAGAAACAGACAGCTCCTCCCTTGAGAAATCCCTGTGATTT 1322  
|||||  
Db 1 CCCAATGTGAGGGCGAGAAACAGACAGCTCCTCCCTTGAGAAATCCCTGTGATTT 60

RESULT 13  
ABN58591  
ID ABN58591 standard; DNA; 60 BP.

AC ABN58591;  
XX  
DT 15-JUL-2002 (first entry)

DE Human spliced transcript detection oligonucleotide SEQ ID NO:31339.

KW Human; mouse; rat; splice transcript; detection; RNA transcript;  
KW splice variant; transcriptome; oligonucleotide library; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200210449-A2.  
XX  
PD 07-FEB-2002.  
XX  
PE 20-JUL-2001; 2001WO-IB01903.  
XX  
PR 28-JUL-2000; 2000US-221607P.  
PR 02-MAY-2001; 2001US-287724P.  
XX  
PA (COMP-) COMPUGEN INC.  
XX  
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
XX  
DR WPI; 2002-257383/30.  
XX

PT New oligonucleotide libraries comprising oligonucleotides which  
PT selectively hybridize to mRNAs transcribed from a transcription unit of  
PT a genome, useful for detecting tissue-, pathology-, and  
PT developmental-specific genes

PS Example 1; SEQ ID 31339; 47pp; English.

XX The present invention describes oligonucleotide libraries for detecting  
CC messenger RNAs that populate a (sub-)transcriptome, where the  
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
CC transcription units that populate a genome. The library comprises  
CC several oligonucleotides, each capable of hybridizing selectively to a  
CC set of messenger RNAs transcribed from a given transcription unit of  
CC the genome, which encodes one or more messenger RNA splice variants.  
CC The oligonucleotide libraries are useful for detecting mRNAs from a  
CC biological sample, in expression profiling studies, in qualitatively or  
CC quantitatively characterizing the corresponding transcriptome, and in  
CC detecting RNA transcripts and splice variants of human or animal  
CC transcriptomes. The libraries may also be used as specialised mini  
CC libraries to detect transcripts of a sub-transcriptome under a  
CC particular biological or pathological state, and so allowing the  
CC detection of tissue- and pathology-specific genes such as those genes  
CC only expressed in specific tissue under a specific pathological  
CC condition; to detect developmental specific genes; and to detect RNA  
CC transcripts and splice variants of a transcriptome of a patient suffering  
CC from a particular disorder. ABN27253 to ABN55589 represent  
CC oligonucleotide sequences from rats, humans and mice, which are used in  
CC the exemplification of the present invention.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

Sequence 60 BP; 17 A; 15 C; 15 G; 13 T; 0 other;

Query Match 4.4%; Score 60; DB 24; Length 60;  
Best Local Similarity 100.0%; Pred. No. 8.6e-18;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1263 CCCAAATGTGAGGGCGGCAAGACAGACAGCTCTCCCTTGAGAAATTCCTGTGGATT 1322  
DB 1 CCCAAATGTGAGGGCGGCAAGACAGACAGCTCTCCCTTGAGAAATTCCTGTGGATT 60

RESULT 14  
ABN58593  
ID ABN58593 standard; DNA; 60 BP.  
XX  
XX ABN58593;  
XX  
XX 15-JUL-2002 (first entry)  
XX  
XX Human spliced transcript detection oligonucleotide SEQ ID NO:31341.  
XX  
XX Human; mouse; rat; splice transcript; detection; RNA transcript;  
XX splice variant; transcriptome; oligonucleotide library; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200210449-A2.  
XX  
XX 07-FEB-2002.  
XX  
XX 20-JUL-2001; 2001WO-1B01903.  
XX  
XX 28-JUL-2000; 2000US-221607P.  
XX 02-MAY-2001; 2001US-287724P.  
XX  
XX (COMP-) COMPUGEN INC.  
XX  
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
XX WPI; 2002-257383/30.  
XX  
XX New oligonucleotide libraries comprising oligonucleotides which  
XX selectively hybridize to mRNAs transcribed from a transcription unit of  
XX a genome, useful for detecting tissue-, pathology-, and  
XX developmental-specific genes  
XX  
XX Example 1; SEQ ID 31341; 47pp; English.  
XX  
XX The present invention describes oligonucleotide libraries for detecting  
XX messenger RNAs that populate a (sub-)transcriptome, where the  
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
XX transcription units that populate a genome. The library comprises  
XX several oligonucleotides, each capable of hybridizing selectively to a  
XX set of messenger RNAs transcribed from a given transcription unit of  
XX the genome, which encodes one or more messenger RNA splice variants.  
XX The oligonucleotide libraries are useful for detecting mRNAs from a  
XX biological sample, in expression profiling studies, in qualitatively or  
XX quantitatively characterizing the corresponding transcriptome, and in  
XX detecting RNA transcripts and splice variants of human or animal  
XX transcriptomes. The libraries may also be used as specialised mini  
XX libraries to detect transcripts of a sub-transcriptome under a  
XX particular biological or pathological state, and so allowing the  
XX detection of tissue- and pathology-specific genes such as those genes  
XX only expressed in specific tissue under a specific pathological  
XX condition; to detect developmental specific genes; and to detect RNA  
XX transcripts and splice variants of a transcriptome of a patient suffering  
XX from a particular disorder. ABN27253 to ABN59589 represent  
XX oligonucleotide sequences from rats, humans and mice, which are used in  
XX the exemplification of the present invention.  
XX N.B. The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pcl\_sequences.

Sequence 60 BP; 12 A; 19 C; 10 G; 19 T; 0 other;

Query Match 4.4%; Score 60; DB 24; Length 60;  
Best Local Similarity 100.0%; Pred. No. 8.6e-18;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

979 GGCTCCTGGGCGCTTCACGTTTCCATCCCAATTAATACAGTATTCACACTCTTA 1038  
DB 1 GGCTCCTGGGCGCTTCACGTTTCCATCCCAATTAATACAGTATTCACACTCTTA 60

RESULT 15  
ABN58848  
ID ABN58848 standard; DNA; 60 BP.  
XX  
XX ABN58848;  
XX  
XX 15-JUL-2002 (first entry)  
XX  
XX Human spliced transcript detection oligonucleotide SEQ ID NO:31596.  
XX  
XX Human; mouse; rat; splice transcript; detection; RNA transcript;  
XX splice variant; transcriptome; oligonucleotide library; ss.  
XX  
XX Homo sapiens.  
XX  
XX WO200210449-A2.  
XX  
XX 07-FEB-2002.  
XX  
XX 20-JUL-2001; 2001WO-1B01903.  
XX  
XX 28-JUL-2000; 2000US-221607P.  
XX 02-MAY-2001; 2001US-287724P.  
XX  
XX (COMP-) COMPUGEN INC.  
XX  
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;  
XX WPI; 2002-257383/30.  
XX  
XX New oligonucleotide libraries comprising oligonucleotides which  
XX selectively hybridize to mRNAs transcribed from a transcription unit of  
XX a genome, useful for detecting tissue-, pathology-, and  
XX developmental-specific genes  
XX  
XX Example 1; SEQ ID 31596; 47pp; English.  
XX  
XX The present invention describes oligonucleotide libraries for detecting  
XX messenger RNAs that populate a (sub-)transcriptome, where the  
XX (sub-)transcriptome comprises messenger RNAs transcribed from multiple  
XX transcription units that populate a genome. The library comprises  
XX several oligonucleotides, each capable of hybridizing selectively to a  
XX set of messenger RNAs transcribed from a given transcription unit of  
XX the genome, which encodes one or more messenger RNA splice variants.  
XX The oligonucleotide libraries are useful for detecting mRNAs from a  
XX biological sample, in expression profiling studies, in qualitatively or  
XX quantitatively characterizing the corresponding transcriptome, and in  
XX detecting RNA transcripts and splice variants of human or animal  
XX transcriptomes. The libraries may also be used as specialised mini  
XX libraries to detect transcripts of a sub-transcriptome under a  
XX particular biological or pathological state, and so allowing the  
XX detection of tissue- and pathology-specific genes such as those genes  
XX only expressed in specific tissue under a specific pathological  
XX condition; to detect developmental specific genes; and to detect RNA  
XX transcripts and splice variants of a transcriptome of a patient suffering  
XX from a particular disorder. ABN27253 to ABN59589 represent  
XX oligonucleotide sequences from rats, humans and mice, which are used in  
XX the exemplification of the present invention.  
XX N.B. The sequence data for this patent did not form part of the printed  
XX specification, but was obtained in electronic format directly from WIPO  
XX at ftp.wipo.int/pub/published\_pcl\_sequences.

Sequence 60 BP; 13 A; 16 C; 17 G; 14 T; 0 other;

Query Match 4.4%; Score 60; DB 24; Length 60;

Best Local Similarity 100.0%; Pred. No. 8.6e-18;

Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1145 GCATTGTGTTCACTGTCTGTGGCAAGATGGGTCCAGAAGACCCCACTTCAGGCAC 1204

DB 1 GCATTGTGTTCACTGTCTGTGGCAAGATGGGTCCAGAAGACCCCACTTCAGGCAC 60

Search completed: May 8, 2003, 12:20:10  
Job time : 347 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 10:21:50 ; Search time 3521 Seconds

(without alignments)  
11348.524 Million cell updates/sec

Title: US-09-245-198a-3

Perfect score: 1373  
Sequence: 1 atgcattgttagcttga.....gacaaatgttgaatg 1373

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size: 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database:

GenDb1:  
1: gb\_da:\*  
2: gb\_hlg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hlg\_hum:\*  
31: em\_hlg\_inv:\*  
32: em\_hlg\_other:\*  
33: em\_hlg\_mus:\*  
34: em\_hlg\_pln:\*  
35: em\_hlg\_rtd:\*  
36: em\_hlg\_mam:\*  
37: em\_hlg\_vrl:\*  
38: em\_sy:\*  
39: em\_hlgo\_hum:\*  
40: em\_hlgo\_mus:\*  
41: em\_hlgo\_other:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1285	93.6	1306	9 AF030099	AF030099 Homo sapi
2	1172	85.4	1353	6 AX201324	AX201324 Sequence
3	1172	85.4	1368	9 AF055872	AF055872 Homo sapi
4	958	69.8	1236	6 AR140407	AR140407 Sequence
5	796	58.0	1651	9 BC019047	BC019047 Homo sapi
6	625	45.5	898	6 AX180714	AX180714 Sequence
7	620	45.2	177703	2 AC016876	AC016876 Homo sapi
8	435	31.7	215795	2 AC127470	AC127470 Pan trogl
9	179	13.0	195	6 AX379024	AX379024 Sequence
10	145	10.6	177555	2 AC130192	AC130192 Sus scrof
11	119	8.7	138792	2 AC119115	AC119115 Rattus no
12	119	8.7	203083	2 AC069459	AC069459 Mus muscu
13	119	8.7	234182	10 AL603707	AL603707 Mouse DNA
14	59	4.3	184026	2 AC098923	AC098923 Rattus no
15	58	4.2	135533	2 AC126921	AC126921 Bos tauru
16	58	4.2	161428	2 AC126925	AC126925 Canis fam
17	50	3.6	50	6 AX201395	AX201395 Sequence
18	50	3.6	203281	2 AC126237	AC126237 Canis fam
19	48	3.5	164504	2 AC129071	AC129071 Pan trogl
20	46	3.4	1168	10 AF030100	AF030100 Mus muscu
21	34	2.5	139405	2 AC126239	AC126239 Felis cat
22	26	1.9	26	6 AX201459	AX201459 Sequence
23	26	1.9	153	6 AX322295	AX322295 Sequence
24	26	1.9	281	6 AX209050	AX209050 Sequence
25	26	1.9	391	6 AX198506	AX198506 Sequence
26	26	1.9	418	9 BC032480	BC032480 Homo sapi
27	26	1.9	439	6 AX330518	AX330518 Sequence
28	26	1.9	439	6 AX410097	AX410097 Sequence
29	26	1.9	483	6 AX381620	AX381620 Sequence
30	26	1.9	487	6 AX408082	AX408082 Sequence
31	26	1.9	531	6 AX381939	AX381939 Sequence
32	26	1.9	2183	9 HSM80267	HSM80267 Homo sapi
33	26	1.9	110293	9 HSDJ6863	AL049766 Homo sapi
34	26	1.9	159020	2 AC027023	AC027023 Homo sapi
35	26	1.9	203281	2 AC126237	AC126237 Canis fam
36	23	1.7	23	6 AX201458	AX201458 Sequence
37	23	1.7	160340	2 AL845513	AL845513 Danio rer
38	23	1.7	185279	2 AC093996	AC093996 Rattus no
39	22	1.6	1173	1 AF182516	AF182516 Pseudom
40	22	1.6	2942	10 MM0250723	AJ250723 Mus muscu
41	22	1.6	10989	10 AF45018S4	AF450141 Mus muscu
42	22	1.6	41037	2 AC017476	AC017476 Drosophi
43	22	1.6	49536	2 AC130967	AC130967 Rattus no
44	22	1.6	58107	2 AC106343	AC106343 Rattus no
45	22	1.6	62952	2 AC111786	AC111786 Rattus no

# ALIGNMENTS

RESULT 1	AF030099	1306 bp	mrna	linear	PRI 20-DEC-1997
LOCUS	AF030099				
DEFINITION	Homo sapiens TWEAK mRNA, complete cds.				
ACCESSION	AF030099				
VERSION	AF030099.1	GI:2707218			
KEYWORDS					
SOURCE					
ORGANISM	Homo sapiens.				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE	1 (bases 1 to 1306)				
AUTHORS	Chicheportliche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.,				
	Hession,C., Garcia,I. and Browning,J.L.				
TITLE	TWEAK, a new secreted ligand in the tumor necrosis factor family				









```

QY 429 AGCCATTATGAGTTCATCCAGACCTGAGACGAGGAGCGCAGGACGAGTGTGACG 488
    |||||||
Db 324 AGCCATTATGAGTTCATCCAGACCTGAGACGAGGAGCGCAGGACGAGTGTGACG 383
QY 489 GACAGGATGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 548
    |||||||
Db 384 GACAGGATGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 443
QY 549 CCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 608
    |||||||
Db 444 CCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 503
QY 609 CTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 668
    |||||||
Db 504 CTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 563
QY 669 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 728
    |||||||
Db 564 CCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 623
QY 729 CCTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 788
    |||||||
Db 624 CCTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 683
QY 789 CCTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 848
    |||||||
Db 684 CCTGTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 743
QY 849 TCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 908
    |||||||
Db 744 TCAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 803
QY 909 CTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 968
    |||||||
Db 804 CTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 863
QY 969 TCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1028
    |||||||
Db 864 TCCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 923
QY 1029 CCCACTTATCTTACACGCTCCCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1088
    |||||||
Db 924 CCCACTTATCTTACACGCTCCCGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 983
QY 1089 CTGACCCCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1148
    |||||||
Db 984 CTGACCCCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1043
QY 1149 TGTGTTCACTGTAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1208
    |||||||
Db 1044 TGTGTTCACTGTAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1103
QY 1209 AGGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1268
    |||||||
Db 1104 AGGGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1163
QY 1269 TGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1328
    |||||||
Db 1164 TGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1223
QY 1329 CAGATATATATTTT 1341
    |||||||
Db 1224 CAGATATATATTTT 1236
    |||||||

```

RESULT 5

LOCUS

DEFINITION

ACCESSION

VERSION

1651 bp mRNA linear PRI 11-DEC-2001

BC019047

Home sapiens, Similar to tumor necrosis factor (ligand)

BC019047

GI:17512138

KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

REMARK  
COMMENT

MGc.  
Homo sapiens.  
Homo sapiens.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 1651)  
Strausberg, R.  
Direct Submission  
Submitted (07-DEC-2001) National Institutes of Health, Mammalian  
Gene Collection (MGc), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
Contact: MGC help desk  
Email: [cgaphs-remail.nih.gov](mailto:cgaphs-remail.nih.gov)  
Tissue Procurement: Louis Staudt  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNI)  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: <http://www.nisc.nih.gov/>  
Contact: [nisc\\_mgc@nih.gov](mailto:nisc_mgc@nih.gov)

Shenchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,  
Benjamin, B., Blakesley, R.W., Bouffard, G., Brinkley, C., Brooks, S.,  
Dierlich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
Lim, M., Maduro, O.L., Masello, C., Mastrian, S.D., McCloskey, J.C.,  
McDowell, J., Pearson, R., Snyder, B., Statilpop, S., Thomas, P.J.,  
Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,  
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/ILNI at: <http://image.llnl.gov>  
Series: IRIL Plate: 30 Row: P Column: 5  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 4507596.

FEATURES

source

1. 1651

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="MGC:20669 IMAGE:4766071"

/tissue\_type="Primary B-cells from Tonsils"

/clone\_11b="NIH-MGC\_48"

/lab\_host="DH10B-R"

/note="Vector: pOTB7"

106. 510

/codon\_start=1

/product="Similar to tumor necrosis factor (ligand)

superfamily, member 12"

/protein\_id="AAH19047.1"

/db\_xref="GI:17512139"

/translation="MAARRSRRRRRREPGTALLVPLAGLGLAAGLILAVSL  
GSRASLSQEPRAQELVAEDDDSELPNPOBESDPPFLNRLVPRPSAPKGRKR  
ARRAIAHTEVHPRPDGAQADQDGYTCLRP"

BASE COUNT

344 a 517 c 481 g 309 t

ORIGIN

Query Match 58.0%; Score 796; DB 9; Length 1651;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 896; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 476 CAGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 535

Db 715 CAGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 774

QY 536 TCGGCTACACCGCCAGATCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 595

Db 775 TCGGCTACACCGCCAGATCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 834

QY 596 ACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 655

Db 835 ACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 894





RESULT 8  
 AC127470  
 LOCUS  
 DEFINITION Pan troglodytes clone RP43-145D13, WORKING DRAFT SEQUENCE, 15  
 unorderd pieces.  
 AC127470  
 AC127470.1 GI:2186866  
 HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE  
 ORGANISM  
 Pan troglodytes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.  
 REFERENCE  
 AUTHORS  
 Aktter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,  
 Benjann, B., Blakesley, R.W., Bouffard, G., Breen, K., Brinkley, C.,  
 Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,  
 Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Latic, P.,  
 Lee-Lin, S.-O., Legaspi, R., Maduro, O.L., Maduro, V.B.,  
 Margulies, E.H., Mastello, C., Maskeri, B., Mastrian, S.D.,  
 McCloskey, J.C., McDowell, J., Paguligan, C., Pearson, R.,  
 Portnoy, M.E., Prasad, A., Schueler, M.G., Stantlipo, S., Thomas, J.W.,  
 Thomas, P.J., Touchman, J.W., Tsurgren, C., Vogt, J.L., Walker, M.A.,  
 Weherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.  
 NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 215795)  
 Green, E.D.  
 Direct Submission  
 Submitted (17-JUL-2002) NIH Intramural Sequencing Center, 8717  
 Government Circle, Gaithersburg, MD 20877, USA  
 COMMENT  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 JOURNAL  
 TITLE  
 JOURNAL  
 CENTER: NIH Intramural Sequencing Center  
 Web site: <http://www.nisc.nih.gov>  
 Contact: [nisc-zoo@nhgri.nih.gov](mailto:nisc-zoo@nhgri.nih.gov)  
 Project Information  
 Center project name: cms  
 Center clone name: 145D13  
 Summary Statistics  
 Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 203929 bases at least Q40  
 Consensus quality: 206865 bases at least Q20  
 Consensus quality: 209181 bases at least Q30  
 Insert size: 19000; agarose-1p  
 Insert size: 214395; sum-of-coverage  
 Quality coverage: 7.30x in Q20 bases; agarose-1p  
 Quality coverage: 6.47x in Q20 bases; sum-of-coverage  
 NOTE: This is a 'working draft' sequence. It currently  
 consists of 15 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.  
 1 3593: contig of 3593 bp in length  
 3594 3693: gap of unknown length  
 3694 8001: contig of 4308 bp in length  
 8002 8101: gap of unknown length  
 8102 15639: contig of 7538 bp in length  
 15640 15739: gap of unknown length  
 15740 24775: contig of 9036 bp in length  
 24776 24875: gap of unknown length  
 24876 30150: contig of 5275 bp in length  
 30151 30250: gap of unknown length  
 30251 39212: contig of 8962 bp in length  
 39213 39312: gap of unknown length  
 39313 49342: contig of 10030 bp in length  
 49343 49442: gap of unknown length

FEATURES  
 source  
 49443 60956: contig of 11514 bp in length  
 60957 61056: gap of unknown length  
 61057 74520: contig of 13464 bp in length  
 74521 74620: gap of unknown length  
 74621 89522: contig of 14902 bp in length  
 89523 89622: gap of unknown length  
 89623 105537: contig of 15915 bp in length  
 105538 121197: gap of unknown length  
 121198 121297: contig of 15560 bp in length  
 121298 138883: contig of 17586 bp in length  
 138884 138983: gap of unknown length  
 138984 168802: contig of 27819 bp in length  
 168803 215795: gap of unknown length  
 16903 215795: contig of 48893 bp in length.  
 Location/Qualifiers  
 1. 215795  
 /organism="Pan troglodytes"  
 /db\_xref="taxon:9598"  
 /clone="RP43-145D13"  
 /clone\_1ID="RP43"  
 1. 3593  
 /note="assembly-fragment"  
 3694. 8001  
 /note="assembly-fragment"  
 8102. 15639  
 /note="assembly-fragment"  
 15740. 24775  
 /note="assembly-fragment"  
 24876. 30150  
 /note="assembly-fragment"  
 30251. 39212  
 /note="assembly-fragment"  
 39313. 49342  
 /note="assembly-fragment"  
 49443. 60956  
 /note="assembly-fragment"  
 61057. 74520  
 /note="assembly-fragment"  
 74621. 89522  
 /note="assembly-fragment"  
 89623. 105537  
 /note="assembly-fragment"  
 105638. 121197  
 /note="assembly-fragment"  
 121298. 138883  
 /note="assembly-fragment"  
 138984. 168802  
 /note="assembly-fragment"  
 16903. 215795  
 /note="assembly-fragment"  
 BASE COUNT 52160 a 54859 c 53752 g 53513 t 1511 others  
 ORIGIN  
 Query Match 31.7%; Score 435; DB 2; Length 215795;  
 Best Local Similarity 99.3%; Pred. No. 6.2e-232;  
 Matches 755; Conservative 0; Mismatches 4; Indels 1; Gaps 1;  
 QY 601 CAGGTGACCTTGTATGAGGGAAGGCTGTCTACCTGAAGCTGAGCTGTGTGATGATG  
 Db 199456 CAGGTGACCTTGTATGAGGGAAGGCTGTCTACCTGAAGCTGAGCTGTGTGATGATG  
 QY 661 GTGCTGGCCCTGCGCTGCTGAGGAATCTCAGCCACTGCGGCGAGTTCCTCGGGCCC  
 Db 199516 GTGCTGGCCCTGCGCTGCTGAGGAATCTCAGCCACTGCGGCGAGTTCCTCGGGCCC  
 QY 721 CAGTTCGCTTCGCGGAGGTGTCTGAGCTTTGGCCCTGCGGCGAGGCTCTCCGCGG  
 Db 199576 CAGTTCGCTTCGCGGAGGTGTCTGAGCTTTGGCCCTGCGGCGAGGCTCTCCGCGG  
 QY 781 ATCCGACCTCCCGCGGCGGCAATCAAGCGCCCGCTTCTCACTCACTCGGACATC  
 Db 199636 ATCCGACCTCCCGCGGCGGCAATCAAGCGCCCGCTTCTCACTCACTCGGACATC

```

OY 841 TTCAGAGTTCAGTGAAGGGGCTGGTCTCCACAGTCTGCCAGAGCTGGGCTCCCT
    |||||||
Db 199696 TTCAGAGTTCAGTGAAGGGGCTGGTCTCCACAGTCTGCCAGAGCTGGGCTCCCT 199755
OY 901 GCACAGCTCTGTGGGACCCGGTCCCTGTGCCACACCTTCAGCGCTTTGTCTCAGA
    |||||||
Db 199756 GCACAGCTCTGTGGGACCCGGTCTGTGCCACACCTTCAGCGCTTTGTCTCAGA 199815
OY 961 CCTGGCCCTCTCTCTAGAGGCTGCTGGGCTGTCTCAGTGTTCATCCACATTAAT 1020
    |||||||
Db 199816 CCTGGCCCTCTCTCTAGAGGCTGCTGGGCTGTCTCAGTGTTCATCCACATTAAT 199875
OY 1021 ACAGATTTCCCACTCTTATCTTACAACT-CCCCACCGCCCACTGCCACTCAGT 1079
    |||||||
Db 199876 ACAGATTTCCCACTCTTATCTTACAACTCCCCACCGCCCACTGCCACTCAGT 199935
OY 1080 CCCCAGTCCCTGACCCCTTTGAGGCCCCCAGTATCTGCATCCCTCCCTGGCCACAGACC 1139
    |||||||
Db 199936 CCCCAGTCCCTGACCCCTTTGAGGCCCCCAGTATCTGCATCCCTCCCTGGCCACAGACC 199995
OY 1140 CCAGGCGATTTGTTCTACTCTGTCTGTGGCAAGATGGTTCAGAGACCCCACTTAA 1199
    |||||||
Db 199996 CCAGGCGATTTGTTCTACTCTGTCTGTGGCAAGATGGTTCAGAGACCCCACTTAA 200055
OY 1200 GGCACCTAAGAGGGGCTGACCTGGCGGCGAGAACCAAGAGAGAGAGGCGCTAGGCCAGGA 1259
    |||||||
Db 200056 GGCACCTAAGAGGGGCTGACCTGGCGGCGAGAACCAAGAGAGAGAGGCGCTAGGCCAGGA 200115
OY 1260 GTTCCCAATGTGAGGGGGGAGAAACAGACAAAGCTCTCCCTTGAGATTCCTGTGGA 1319
    |||||||
Db 200116 GTTCCCAATGTGAGGGGGGAGAAACAGACAAAGCTCTCCCTTGAGATTCCTGTGGA 200175
OY 1320 TTTTAAAAACATATATTTATTTATTTATTTGTGACAAA 1359
    |||||||
Db 200176 TTTTAAAAACATATATTTATTTATTTATTTGTGACAAA 200215

RESULT 9
AX379024 195 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 66 from Patent WO0196389.
ACCESSION AX379024
VERSION AX379024.1 GI:19574864
KEYWORDS
SOURCE
ORGANISM human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Meagher, M.J., King, G.E., Xu, J. and Secrist, H.
COMPOSITIONS and methods for the therapy and diagnosis of colon
cancer.
PATENT: WO 0196389-A 66 20-DEC-2001;
JOURNAL CORIXA CORPORATION (US)
FEATURES
source 1. 195
Location/Qualifiers
BASE COUNT 49 a 51 c 58 g 37 t
ORIGIN
Query Match 13.0%; Score 179; DB 6; Length 195;
Best Local Similarity 100.0%; Pred. NO. 1.6e-88;
Matches 179; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1146 CATGTGTTCACTGACTCTGTGGCAAGATGGTTCAGAGACCCCACTTCAGGCACT 1205
    |||||||
Db 17 CATGTGTTCACTGACTCTGTGGCAAGATGGTTCAGAGACCCCACTTCAGGCACT 76
OY 1206 AAGAGGGCTGACCTGGCGGCGAGAGCAAGAGACTGGCTTAGGCCAGAGCTTCCC 1265
    |||||||
Db 77 AAGAGGGCTGACCTGGCGGCGAGAGCAAGAGACTGGCTTAGGCCAGAGAGTCCC 136

```

```

OY 1266 AATGTGAGGGGCGAGAAACAGACAGCTCTCCCTTGAGATTCCTGTGATTTT 1324
    |||||||
Db 137 AATGTGAGGGGCGAGAAACAGACAGCTCTCCCTTGAGATTCCTGTGATTTT 195

RESULT 10
AC130192/c
LOCUS
DEFINITION
AC130192 177555 bp DNA linear HTG 08-AUG-2002
Sus scrofa clone RP44-436K21, WORKING DRAFT SEQUENCE, 10 unordered
pieces.
AC130192
AC130192.1 GI:22138443
HTG: HTGS_PHASE1, HTGS_DRAFT.
SOURCE
pig.
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 177555).
Akther, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,
Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J.,
Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P.,
Lee-Ihn, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B.,
Margulies, E.H., Masello, C., Maskeri, B., Mastrian, S.D.,
McCluskey, J.C., McDowell, J., Paguirigan, C., Pearson, R.,
Portnoy, M.E., Prasad, A., Schueler, M.G., Stantirip, S., Thomas, J.W.,
Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, J.L., Walker, M.A.,
Wetherby, R.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
NISC Comparative Sequencing Initiative.
Unpublished
2 (bases 1 to 177555)
Green, E.D.
Direct Submission
Submitted (08-AUG-2002) NIH Intramural Sequencing Center, 8717
Grovefront Circle, Gaithersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoehngri.nih.gov
----- Project Information
Center project name: dbg
Center clone name: 436K21
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 164120 bases at least Q40
Consensus quality: 168085 bases at least Q30
Consensus quality: 170767 bases at least Q20
Insert size: 169000; agarose-fp
Insert size: 17655; sum-of-coverage
Quality coverage: 4.78x in Q20 bases; agarose-fp
Quality coverage: 4.57x in Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 3460: contig of 3460 bp in length
* 3461 3560: gap of unknown length
* 3561 6869: contig of 3309 bp in length
* 6870 6969: gap of unknown length
* 6970 13465: contig of 6496 bp in length
* 13466 13565: gap of unknown length
* 13566 21301: contig of 7736 bp in length
* 21302 21402: gap of unknown length
* 21402 33719: contig of 12318 bp in length

```



(see [http://www.hgsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 32 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```
1 1022: contig of 1022 bp in length
* 1023 1122: gap of unknown length
* 1123 2394: contig of 1272 bp in length
* 2395 2494: gap of unknown length
* 2495 3663: contig of 1169 bp in length
* 3664 3763: gap of unknown length
* 3764 4879: contig of 1116 bp in length
* 4880 4979: gap of unknown length
* 4980 6787: contig of 1808 bp in length
* 6788 6888: gap of unknown length
* 6889 8192: contig of 1305 bp in length
* 8193 8292: gap of unknown length
* 8293 9461: contig of 1169 bp in length
* 9462 9561: gap of unknown length
* 9562 11435: contig of 1874 bp in length
* 11436 11535: gap of unknown length
* 11536 13535: contig of 2000 bp in length
* 13536 13635: gap of unknown length
* 13636 17323: contig of 3688 bp in length
* 17324 17424 20406: contig of 2983 bp in length
* 20407 20507 23445: contig of 2939 bp in length
* 23446 23545: gap of unknown length
* 23546 25600: contig of 2055 bp in length
* 25601 25701 28049: contig of 2349 bp in length
* 28050 28149: gap of unknown length
* 28150 30474: contig of 2325 bp in length
* 30475 30574: gap of unknown length
* 30575 35498: contig of 4924 bp in length
* 35499 35598: gap of unknown length
* 35599 39284: contig of 3686 bp in length
* 39285 39385 41970: gap of unknown length
* 39386 42070: contig of 2586 bp in length
* 42071 45659: gap of unknown length
* 45660 45759: gap of unknown length
* 45760 50332: contig of 4573 bp in length
* 50333 50432: gap of unknown length
* 50433 54695: contig of 4263 bp in length
* 54696 54795: gap of unknown length
* 54796 58289: contig of 3494 bp in length
* 58290 58390 63599: contig of 5210 bp in length
* 63600 63699: gap of unknown length
* 63700 70305: contig of 6606 bp in length
* 70306 70405: gap of unknown length
* 70406 76123: contig of 5718 bp in length
* 76124 76223: gap of unknown length
* 76224 84961: contig of 8738 bp in length
* 84962 85061: gap of unknown length
* 85062 93614: contig of 8553 bp in length
* 93615 93714: gap of unknown length
* 93715 103352: contig of 9638 bp in length
* 103353 103452: gap of unknown length
* 103453 110299: contig of 6847 bp in length
* 110300 110399: gap of unknown length
* 116180: contig of 5781 bp in length
* 116181 116280: gap of unknown length
* 116281 125484: contig of 9204 bp in length
* 125485 125584: gap of unknown length
* 125585 138792: contig of 13208 bp in length.
```

FEATURES

Location/Qualifiers

source

1. 138792

/organism="Rattus norvegicus"

/db\_xref="taxon:10116"

/clone="CH230-320N23"

BASE COUNT 34391 a 31996 c 31712 g 36193 t 4500 others

ORIGIN

Query Match

8.7%; Score 119; DB 2: Length 138792;

Best Local Similarity 100.0%; Pred. No. 1,4e-34;

Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 51 ATCCCTCGGTCGCGGATGGGGGCGGTGAGGAGGACAGCCCGCCCGCCATGCG 110

Db 129318 ATCCCTCGGTCGCGGATGGGGGCGGTGAGGAGGACAGCCCGCCCGCCATGCG 129377

Oy 111 CGCCGTCGAGCCAGAGCGGAGGGCGCCGGGGAGCCCGGACCCCTGCTGG 169

Db 129378 CGCCGTCGAGCCAGAGCGGAGGGCGCCGGGGAGCCCGGACCCCTGCTGG 129436

RESULT 12

AC069459/c

AC069459 203083 bp DNA linear HTG 27-JUN-2001

LOCUS Mus musculus chromosome 11 clone RP23-168P5, WORKING DRAFT

DEFINITION SEQUENCE, 7 unordered pieces.

AC069459

AC069459.23 GI:14547768

KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.

VERSION

REWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

Estimated insert size: 21056; sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-gel estimation  
Quality coverage: 7.2x in Q20 bases; sum-of-contigs estimation

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.bjsg.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.bjsg.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 62152: contig of 62152 bp in length  
\* 62153 62252: gap of unknown length  
\* 62253 118772: contig of 56520 bp in length  
\* 118773 118872: gap of unknown length  
\* 118873 148924: contig of 30052 bp in length  
\* 148925 149024: gap of unknown length  
\* 149025 167331: contig of 18207 bp in length  
\* 167332 189907: gap of unknown length  
\* 189908 190007: contig of 22576 bp in length  
\* 190008 196537: gap of unknown length  
\* 196538 203083: contig of 6446 bp in length.  
\* Location/Qualifiers  
1. 203083

FEATURES  
Source  
1. 203083  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="11"  
/clone="RP23-168P5"

BASE COUNT 52662 a 49293 c 47892 g 52633 t 603 others  
ORIGIN

Query Match 8.7%; Score 119; DB 2; Length 203083;  
Best Local Similarity 100.0%; Pred. No. 1.4e-54;  
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ATCCCTGGGTCGCCGATGGGGGGCGGTGAGGACAGACAGCCCCCGCCATGCG 110  
|||||  
Db 51733 ATCCCTGGGTCGCCGATGGGGGGCGGTGAGGACAGACAGCCCCCGCCATGCG 51674

QY 111 CGCCCGTGGAGCGAGAGGGGCGCCCGGGGAGACCGCCCTGCTGG 169  
|||||  
Db 51673 CGCCCGTGGAGCGAGAGGGGCGCCCGGGGAGACCGCCCTGCTGG 51615

RESULT 13  
AL603707/c 234182 bp DNA linear ROD 17-NOV-2001  
LOCUS  
DEFINITION Mouse DNA sequence from clone RP23-422L16 on chromosome 11,  
complete sequence.  
ACCESSION AL603707  
VERSION AL603707.5 GI:17017790  
KEYWORDS HTG.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 234182)

REFERENCE 1 (bases 1 to 234182)  
AUTHORS Pearce, A.  
TITLE Direct Submissio  
JOURNAL Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humayr@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
On Nov 20, 2001 this sequence version replaced gi:16605765.

COMMENT During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submissio  
corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given  
in the feature table with their source databases: Em: EMBL; Sw:  
SWISSPROT; Tr: TrEMBL; Wp: WormPep; Information on the WormPep  
database can be found at  
<http://www.sanger.ac.uk/projects/C.elegans/wormpep> RP23-422L16 is  
from the RPI-23 Mouse PAC library  
constructed by the group of Pieter de Jong.  
For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pBAC3.6  
This sequence is the entire insert of clone RP23-422L16.

FEATURES  
Source  
1. 234182  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="11"  
/clone="RP23-422L16"  
/clone\_lib="RPI-23"  
/complement="84050..84131"  
/note="Sequence from uni-directional primer reads and dGTP  
big dye terminator reads only."

BASE COUNT 59310 a 56824 c 57519 g 60529 t  
ORIGIN

Query Match 8.7%; Score 119; DB 10; Length 234182;  
Best Local Similarity 100.0%; Pred. No. 1.4e-54;  
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 ATCCCTGGGTCGCCGATGGGGGGCGGTGAGGACAGACAGCCCCCGCCATGCG 110  
|||||  
Db 75038 ATCCCTGGGTCGCCGATGGGGGGCGGTGAGGACAGACAGCCCCCGCCATGCG 74979

QY 111 CGCCCGTGGAGCGAGAGGGGCGCCCGGGGAGACCGCCCTGCTGG 169  
|||||  
Db 74978 CGCCCGTGGAGCGAGAGGGGCGCCCGGGGAGACCGCCCTGCTGG 74920

RESULT 14  
AC098923/c 184026 bp DNA linear HTG 12-JUL-2002  
LOCUS  
DEFINITION Rattus norvegicus clone CH230-154B15, \*\*\* SEQUENCING IN PROGRESS  
ACCESSION AC098923  
VERSION AC098923.5 GI:21729926  
KEYWORDS HTGS PHASE1.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 184026)

REFERENCE 1 (bases 1 to 184026)  
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,  
Alsbrooks, S.L., Amaral, H.C., Aré, J.R., Ayala, M., Banks, T.,  
Barbieri, J., Benton, J., Bimge, K., Blankenburg, K., Bonini, D.,  
Bouck, J., Bowle, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,  
Bukey, C., Burck, P., Burkett, C., Burrell, R.L., Byrd, N.C.,  
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Devilla, M.L., Davis, C., Davy-Carroll, L., Dedrich, D.A.,  
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
Doutwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabisi, A., Gao, J., Garcia, A., Garner, F., Garza, N., Gill, R.,  
Gorrell, J.R., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,



Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,  
Honsi,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Joliver,S., Joudah,S.,  
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,  
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H.,  
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,  
Meheswarl,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,  
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,  
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,  
Mosier,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokkwo,S., Ogun,M., Okunolu,G.,  
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,  
Scheerer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,  
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,  
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,  
Umami,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wleczyk,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,  
Weinstock,G. and Gibbs,R.

Unpublished  
Direct Submission  
2 (bases 1 to 184026)  
Worley,K.C.

Submitted (06-NOV-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 184026)  
Worley,K.C.

Direct Submission  
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 11, 2002 this sequence version replaced gi:17973852.

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: G10K  
Center clone name: CH230-154B15  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 109715 bases at least Q40  
Consensus quality: 115369 bases at least Q30  
Consensus quality: 121073 bases at least Q20  
-----

\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a working draft sequence. It currently  
\* consists of 71 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1385: contig of 1385 bp in length  
\* 1386 1485: gap of unknown length  
\* 1486 2496: contig of 1011 bp in length  
\* 2497 2596: gap of unknown length  
\* 2597 3886: contig of 1290 bp in length  
\* 3887 3986: gap of unknown length  
\* 3987 5552: contig of 1566 bp in length

5553 5652: gap of unknown length  
\* 5653 7125: contig of 1473 bp in length  
\* 7126 7225: gap of unknown length  
\* 7226 8566: contig of 1341 bp in length  
\* 8567 8666: gap of unknown length  
\* 8667 9765: contig of 1098 bp in length  
\* 9765 9864: gap of unknown length  
\* 9864 11285: contig of 1421 bp in length  
\* 11286 11385: gap of unknown length  
\* 11386 12407: contig of 1022 bp in length  
\* 12408 12507: gap of unknown length  
\* 12508 13827: contig of 1320 bp in length  
\* 13828 13927: gap of unknown length  
\* 13928 15227: contig of 1300 bp in length  
\* 15228 15327: gap of unknown length  
\* 15328 16441: contig of 1114 bp in length  
\* 16442 16541: gap of unknown length  
\* 16542 17797: contig of 1256 bp in length  
\* 17798 17897: gap of unknown length  
\* 17898 19717: contig of 1820 bp in length  
\* 19718 19817: gap of unknown length  
\* 19818 21264: contig of 1447 bp in length  
\* 21265 21364: gap of unknown length  
\* 21365 22396: contig of 1032 bp in length  
\* 22397 22496: gap of unknown length  
\* 22497 23618: contig of 1122 bp in length  
\* 23619 23718: gap of unknown length  
\* 23719 25080: contig of 1362 bp in length  
\* 25081 25180: gap of unknown length  
\* 25181 26592: contig of 1412 bp in length  
\* 26593 26692: gap of unknown length  
\* 26693 27883: contig of 1191 bp in length  
\* 27884 27983: gap of unknown length  
\* 27984 29261: gap of unknown length  
\* 29262 29361: contig of 1278 bp in length  
\* 29362 31112: gap of unknown length  
\* 31113 31212: contig of 1751 bp in length  
\* 31213 32801: gap of unknown length  
\* 32802 32901: contig of 1589 bp in length  
\* 32902 34318: gap of unknown length  
\* 34319 34418: contig of 1417 bp in length  
\* 34419 36405: gap of unknown length  
\* 36406 36505: contig of 1987 bp in length  
\* 36506 37644: gap of unknown length  
\* 37645 37744: contig of 1139 bp in length  
\* 37745 39164: gap of unknown length  
\* 39165 39264: contig of 1420 bp in length  
\* 39265 41252: gap of unknown length  
\* 41253 41352: contig of 1988 bp in length  
\* 41353 42492: gap of unknown length  
\* 42493 42592: contig of 1140 bp in length  
\* 42593 44098: gap of unknown length  
\* 44099 44198: contig of 1506 bp in length  
\* 44199 46386: gap of unknown length  
\* 46387 46486: contig of 2188 bp in length  
\* 46487 48067: gap of unknown length  
\* 48068 48167: contig of 1581 bp in length  
\* 48168 49678: gap of unknown length  
\* 49679 49778: contig of 1511 bp in length  
\* 49779 51600: gap of unknown length  
\* 51601 51700: contig of 1822 bp in length  
\* 51701 53618: gap of unknown length  
\* 53619 53718: contig of 1918 bp in length  
\* 53719 56000: gap of unknown length  
\* 56001 56100: contig of 2282 bp in length  
\* 56101 58397: gap of unknown length  
\* 58398 58497: contig of 2297 bp in length  
\* 58498 60619: gap of unknown length  
\* 60619 60718: contig of 2121 bp in length  
\* 60719 62515: gap of unknown length  
\* 62516 65432: contig of 1797 bp in length  
\* 65433 65532: gap of unknown length  
\* 65532: gap of 2817 bp in length

```

* 65533 67184: contig of 1652 bp in length
* 67185 67284: gap of unknown length
* 67285 69655: contig of 2371 bp in length
* 69656 69755: gap of unknown length
* 69756 71983: contig of 2228 bp in length
* 71984 72083: gap of unknown length
* 72084 74212: contig of 2129 bp in length
* 74213 74312: gap of unknown length
* 74313 76381: contig of 2069 bp in length
* 76382 76481: gap of unknown length
* 76482 79698: contig of 3217 bp in length
* 79699 79798: gap of unknown length
* 79799 81766: contig of 1968 bp in length
* 81767 83777: contig of 1911 bp in length
* 83778 83877: gap of unknown length
* 83878 87413: contig of 3535 bp in length
* 87413 87512: gap of unknown length
* 87513 90541: contig of 3029 bp in length
* 90542 90641: gap of unknown length
* 90642 93335: contig of 2694 bp in length
* 93336 93435: gap of unknown length
* 93436 95865: contig of 2430 bp in length
* 95866 95965: gap of unknown length
* 95966 98401: contig of 2436 bp in length
* 98402 98501: gap of unknown length
* 98502 101518: contig of 3017 bp in length

```

```

Query Match      4.3%  Score 59: DB 2: Length 184026;
Best Local Similarity 100.0%: Pred. No. 6.7e-21;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 111 CGCCCGTCGAGCAGAGCGGAGGCGCGCGGAGCGCCGCGCTGCTGG 169
DB 57812 CGCCCGTCGAGCAGAGCGGAGGCGCGCGGAGCGCCGCGCTGCTGG 57754

```

```

RESULT 15
AC126921
LOCUS
DEFINITION
  Bos taurus clone RP42-45D24, WORKING DRAFT SEQUENCE, 12 unordered
  pieces.
ACCESSION
  AC126921
VERSION
  AC126921.1 GI:21724098
KEYWORDS
  HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
  COW.
ORGANISM
  Bos taurus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
  Bovidae; Bovinae; Bos.

```

```

REFERENCE
  1 (bases 1 to 153553)
  Akhtar,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
  Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
  Brooks,S., Dietrich,N.L., Grant,S., Guan,X., Gupta,J.,
  Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Kallins,E., Laric,P.,
  Lee-Lin,S.-Q., Legaspi,R., Maduro,O.L., Maduro,V.B.,
  Margulies,E.H., Mastello,C., Maskeri,B., Mastrian,S.D.,
  McCloskey,J.C., McDowell,J., Pasquigan,C., Pearson,R.,
  Portnoy,M.E., Prasad,A., Schneider,M.G., Stantrilop,S., Thomas,J.W.,
  Thomas,P.J., Touchman,J.W., Tsurgueon,C., Vogt,J.L., Walker,M.A.,
  Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
  NISC Comparative Sequencing Initiative
  Unpublished
  2 (bases 1 to 153553)
  Green,E.D.
  Direct Submission
  Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717
  Grovemont Circle, Gaithersburg, MD 20877, USA
  ----- Genome Center
  Center: NIH Intramural Sequencing Center
  Center code: NISC
  Web site: http://www.nisc.nih.gov
  Contact: nisc_zoehghrl.nih.gov

```

```

COMMENT
  Center: NIH Intramural Sequencing Center
  Center code: NISC
  Web site: http://www.nisc.nih.gov
  Contact: nisc_zoehghrl.nih.gov

```

```

----- Project Information
Center project name: 045D24
Center clone name: ddd1
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 14606 bases at least Q40
Consensus quality: 147748 bases at least Q30
Consensus quality: 148824 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 152453; sum-of-contigs
Quality coverage: 8.80x in Q20 bases; agarose-fp
Quality coverage: 8.72x in Q20 bases; sum-of-contigs
-----

```

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 2312: contig of 2312 bp in length
2313 2412: gap of unknown length
2413 5841: contig of 3429 bp in length
5842 5941: gap of unknown length
5942 8435: contig of 2494 bp in length
8436 8536: gap of unknown length
8536 15789: contig of 7264 bp in length
15789 15899: gap of unknown length
15899 25324: contig of 9335 bp in length
25324 25325: gap of unknown length
25325 32504: contig of 7180 bp in length
32504 32505: gap of unknown length
32505 40970: contig of 8366 bp in length
40970 41070: gap of unknown length
41070 56590: contig of 15520 bp in length
56590 56591: gap of unknown length
56591 73769: contig of 17079 bp in length
73769 73870: gap of unknown length
73870 90860: contig of 16990 bp in length
90860 90959: gap of unknown length
90959 111428: contig of 20469 bp in length
111428 111528: gap of unknown length
111528 111529: contig of 42025 bp in length.

```

## FEATURES

```

source
  1..153553
    /organism="Bos taurus"
    /db_xref="taxon:9913"
    /clone="RP42-45D24"
    /clone_1fb="RP42"
  1..2312
    /note="assembly-fragment"
  2413..5841
    /note="assembly-fragment"
    /note="assembly-fragment"
    vector_side="left"
  5942..8435
    /note="assembly-fragment"
  8536..15789
    /note="assembly-fragment"
  15900..25224
    /note="assembly-fragment"
  25325..32504
    /note="assembly-fragment"
  32603..40970
    /note="assembly-fragment"
  41071..56590
    /note="assembly-fragment"
  56691..73769
    /note="assembly-fragment"
  73870..90859
    /note="assembly-fragment"

```

```
misc_feature      /note="assembly_fragment"  
90960. .111428  
misc_feature      /note="assembly_fragment"  
111529. .153553  
                  /note="assembly_fragment"  
                  clone_end:SP6  
                  vector_side:right"  
BASE COUNT      39178 a 37393 c 36616 g 39259 t 1107 others  
ORIGIN  
Query Match      4.2%; Score 58; DB 2; Length 153553;  
Best Local Similarity 100.0%; Pred. No. 2.4e-20;  
Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 601 CAGTGCACCTTGTGATGAGGGGAAGGCTGTCTACTGAACCTGGACTTCTGCTGGTGATG 658  
      ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
DB 27189 CAGTGCACCTTGTGATGAGGGGAAGGCTGTCTACTGAACCTGGACTTCTGCTGGTGATG 27246
```

Search completed: May 8, 2003, 13:40:41  
Job time : 4822 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 8, 2003, 02:27:27 ; Search time 1056.48 Seconds

(without alignments)  
3449.164 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162

Sequence: 1 VLSGLALACGLLVVSL.....PMHLKAPFLTYRGLFQVH 225

## Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DEV=xlp  
-Q=/qgn2.1/USFto\_spool/US09245198/runat\_06052003.170239.9628/app\_query.fasta.1.846  
-DB-EST-QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=100 -THR\_MAX=100 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09245198 @CGN.1.1.2375 @runat.06052003.170239.9628 -NCPU=6 -ICPU=3  
-NO\_XIPYX -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELop=6 -DELext=7

## Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlin:\*  
5: em\_estlin:\*  
6: em\_estlin:\*  
7: em\_estlin:\*  
8: em\_estlin:\*  
9: gp\_estl:\*  
10: gp\_estl:\*  
11: gp\_estl:\*  
12: gp\_estl:\*  
13: gp\_estl:\*  
14: gp\_estl:\*  
15: em\_estlin:\*  
16: em\_estlin:\*  
17: gp\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_hum:\*  
20: em\_gss\_hum:\*  
21: em\_gss\_hum:\*  
22: em\_gss\_hum:\*  
23: em\_gss\_hum:\*  
24: em\_gss\_hum:\*  
25: em\_gss\_hum:\*  
26: em\_gss\_hum:\*  
27: em\_gss\_hum:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1039	89.4	918	12	BF577781
2	968.5	83.3	731	13	BF577781
3	941.5	81.0	728	13	BF577781
4	882	75.9	777	13	BF577781
5	861	74.1	828	13	BF577781
6	851	74.1	834	13	BF577781
7	857	73.8	861	10	AM765237
8	831.5	71.6	948	14	BQ070185
9	818	70.4	940	14	BQ084231
10	764	65.7	963	14	BQ671259
11	754	64.9	1033	11	AK020909
12	721	62.0	1071	14	BM921213
13	703	60.5	951	14	BQ674188
14	699	60.2	584	10	AM917574
15	670	57.7	445	9	AA870722
16	632.5	54.4	471	9	AA221610
17	629	54.1	650	12	BG404836
18	621	53.4	360	10	BE654876
19	610	52.5	785	13	BF62908
20	561.5	48.3	542	12	BF041509
21	549.5	47.3	404	12	BF044430
22	541	46.6	315	12	BF466521
23	513	44.1	474	13	BF651774
24	506.5	43.6	493	10	BE307031
25	501.5	43.2	894	13	BF908274
26	486	41.8	558	13	BM484863
27	477	41.0	581	13	BF738634
28	473	40.7	538	12	BF821434
29	458	39.4	298	14	BM688946
30	429	36.9	278	10	AM417023
31	428	36.8	342	9	AA637970
32	404.5	34.8	571	12	BF073881
33	404	34.8	345	14	BF57379
34	392	33.7	587	12	BG686319
35	373	32.1	400	9	AI152313
36	373	32.1	441	13	BF967060
37	370	31.8	487	10	AM320117
38	369.5	31.8	698	13	BF906850
39	354	30.5	260	10	BB596688
40	338	29.1	533	10	BB628951
41	336	28.9	910	12	BG110063
42	335	28.8	456	13	BF966255
43	316	27.2	687	14	BQ208433
44	287	24.7	531	13	BF824443
45	243.5	21.0	639	10	BB642326

## ALIGNMENTS

RESULT 1  
LOCUS BF577781 918 bp mRNA linear EST 12-DEC-2000  
DEFINITION 602092080F1 NCI\_CGAP\_C024 Mus musculus cDNA clone IMAGE:4206595 5',  
LOCUS BF577781 mRNA sequence.  
ACCESSION BF577781  
VERSION BF577781.1 GI:11651493  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
REFERENCE 1 (bases 1 to 918)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)



Oy 65 ValArgProArgSerAlaProLysGlyValArgProArgAlaIleAla 84  
 Db 182 GTTGGGCTCCAGAGTGCACCTTAAGCCGGAACACAGCGCTCGAGAGCGATCCCA 241  
 Oy 85 AlaHisTyrGluValHisProArgProGlyInaSpGlyAlaGlnAlaGlyValAspGly 104  
 Db 242 GCCCATATATGAAGTTTCATCCACGACCTGGACAGGAGGCGAGCGAGGTGTGGACGG 301  
 Oy 105 ThrValSerGlyTyrPgluGluThrLysIleAsnSerSerSerProLeuArgTyrAspArg 124  
 Db 302 ACAATGAGTGGCTGGAGAGAGCCAGATCAACAGCTCCACCCCTCGCTACACACCGC 361  
 Oy 125 GlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHis 144  
 Db 362 CAGATCGGGAGATTATATGTACCCGGCTGGCTCTACTACTGTACTGTACAGGTGCAC 421  
 Oy 145 PheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGlyValIleVal 164  
 Db 422 TTGTATGAGGGAGAGGCTGTCTACTGAAGCTGACCTGTGTGTGTGTGTGTGTGTGCC 481  
 Oy 165 LeuArgCysLeuGluGluPheSerAlaThrAlaAspSerProGlyProGlnLeuArg 184  
 Db 482 CTGGCGCTCCGAGAGATTCACACCTCGCCAGCTCCCTCGGGCCCGACGCTCCGC 541  
 Oy 185 LeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArgIleArgThr 204  
 Db 542 CTCGCCAGAGTGTGGGCTGTGGCCCTGGGCGCCAGGTCC-TCCCTCGGATCCGACCC 600  
 Oy 205 LeuProTrrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 224  
 Db 601 CTCCTCGGGCCCATCTCAAGGCTGCCCTTCCTCCACCTACTCGAGACTCTCCAGGTT 660  
 Oy 225 His 225  
 Db 661 CAC 663  
 RESULT 3  
 LOCUS B1870393 728 bp mRNA linear EST 11-OCT-2001  
 DEFINITION 603395641F1 NIH\_MGC\_90 Homo sapiens cDNA clone IMAGE:5405459 5',  
 mRNA sequence.  
 ACCESSION B1870393  
 VERSION B1870393.1 GI:16044066  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 728)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: c9apds-remail.nih.gov  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 plate: L1AM12034 row: a column: 12  
 High quality sequence stop: 728.  
 Location/Qualifiers  
 1..728  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5405459"  
 /clone\_lib="NIH\_MGC\_90"  
 /tissue\_type="adenocarcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: liver; Vector: pCMV-Sport6; Site\_1: NotI;

Site\_2: SalI; Cloned unidirectionally; oligo-dT primed.  
 Average insert size 1.7 kb. Library enriched for  
 full-length clones and constructed by Life Technologies.  
 Note: this is a NIH\_MGC Library."

BASE COUNT 125 a 240 c 227 g 136 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2,99e-89 Length: 728  
 Score: 941.50 Matches: 194  
 Percent Similarity: 91.86% Conservative: 9  
 Best Local Similarity: 87.78% Mismatches: 17  
 Query Match: 81.02% Indels: 3  
 DB: 13 Gaps: 1

US-09-245-198a-2 (1-225) x B1870393 (1-728)

Oy 6 LeuAlaLeuAlaCysLeuGlyLeuLeuLeuValValSerLeuGlySerTrrAlaThr 25  
 Db 2 CTGGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 61  
 Oy 26 LeuSerAla---GlnGluProSerGlnGluGluThrAlaGluAspArgGluPro 44  
 Db 62 CTGTCCGCCAGACAGAGACCTGCCAGAGAGAGCTGTGGCAGAGAGAGACAGACCCG 121  
 Oy 45 ProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPheLeuGluGlnLeu 64  
 Db 122 TCGGACTGAATCCCAACAG 181  
 Oy 65 ValArgProArgSerAlaProLysGlyValArgGlyAlaGlnAlaGlyValAspGly 84  
 Db 182 GTTGGGCTCCAGAGATTCACACCTTAAGCCGGAACACAGCGGCTCGAAGAGCGATCCCA 241  
 Oy 85 AlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGly 104  
 Db 242 GCCCATATATGAAGTTTCATCCACGACCTGGACAGGAGCGAGCGAGGTGTGGACGG 301  
 Oy 105 ThrValSerGlyTyrPgluGluThrLysIleAsnSerSerSerProLeuArgTyrAspArg 124  
 Db 302 ACAATGAGTGGCTGGAGAGAGCCAGATCAACAGCTCCACCCCTCGCTACACACCGC 361  
 Oy 125 GlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHis 144  
 Db 362 CAGATCGGGAGATTATATGTACCCGGCTGGCTCTACTACTGTGTGTGTGTGTGTGTGCC 421  
 Oy 145 PheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGlyValIleVal 164  
 Db 422 TTGTATGAGGGAGAGGCTGTCTACTGAAGCTGACCTGTGTGTGTGTGTGTGTGCC 481  
 Oy 165 LeuArgCysLeuGluGluPheSerAlaThrAlaAspSerProGlyProGlnLeuArg 184  
 Db 482 CTGGCTCCGAGAGATTCACACCTGGGCGCAATCCCTGGGCGCCAGCTCCGC 541  
 Oy 185 LeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArgIleArgThr 204  
 Db 542 CTCGCCAGAGTGTGGGCTGTGGCCCTGGGCGCCAGGTCC-TCCCTCGGATCCGACCC 600  
 Oy 205 LeuProTrrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 224  
 Db 601 CTCCTCGGGCCCATCTCAAGGCTGCCCTTCCTCCACCTACTCGAGACTCTCCAGGTT 659  
 Oy 225 His 225  
 Db 660 CAC 662  
 RESULT 4  
 LOCUS B1819200 777 bp mRNA linear EST 04-OCT-2001  
 DEFINITION 60303461F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5175698 5',  
 mRNA sequence.  
 ACCESSION B1819200  
 VERSION B1819200.1 GI:15930750  
 KEYWORDS EST.





OY 46 GluLeuAsnProGlnThrGluGluSerGlnAspValValProPheLeuGlnGluLeuVal 65  
 DB 44 GAACAGATCCCCACACAGAGAAAGCCAGATCTCGCCCTTCTCTGACCGACTAGTT 103  
 OY 66 ArgProArgArgSerAlaProGlyAlaArgGlyAlaArgProArgArgAlaAlaAla 85  
 DB 104 CGGCTCCACAGAGTGCACCTTAAGCCGGAACACGCGGCTCGAAGACGATCCAGCC 163  
 OY 86 HisTyrGluValHisProArgProGlnGlnAspGlyAlaGlnAlaGlyValAspGlyThr 105  
 DB 164 CATTAATGAAGTTCATCCACGACCTGACAGAGAGCGACAGGTGTGACGAGACA 223  
 OY 106 ValSerGlyTyrPglGluGlnThrLysIleAsnSerSerSerProLeuArgTyrAspArgin 125  
 DB 224 TTGATGGCTGGAGAGAACCCAGAAATCAACAGCTCCAGCCCTGCTGCTACACGCCAG 283  
 OY 126 IleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPhe 145  
 DB 284 ATCGGGAGATTATAGTACCCCGGCTGCTACTACCTGACTGCTGCTGCTGCTGCTGCT 343  
 OY 146 AspGluGlyValAlaValIleuValIleuValIleuValIleuValIleuValIleuVal 165  
 DB 344 GATGAGGGAGAGGCTGTCTACCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 403  
 OY 166 ArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyProGlnLeuArgLeu 185  
 DB 404 CCTCTCTGAGAGAAATTTCTACACCTGCGGCGGAGTCTCTGCGGCGGCGGCGGCGG 463  
 OY 186 CysGlnValSerGlyLeuLeuProLeuArgProGlySerSerSerLeuArgIleArgThrLeu 205  
 DB 464 TGCCAGGTGTCTGGGCTGTGGCCCTGCGGCGGAGGTCCTCCCTGCGAGTCCGACCC 523  
 OY 206 ProTrrAlaHisLeuValAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 225  
 DB 524 CCTGGGCGCATCTCAAGGCTGCCCCCTTCTCACCTTCTGAGCTCTTCAGGTTTCAC 583  
 RESULT 6  
 B1766766 834 bp mRNA linear EST 25-SEP-2001  
 LOCUS 603056866F1 NIH\_MGC\_122 Homo sapiens cDNA clone IMAGE:5206217 5'  
 DEFINITION mRNA sequence.  
 ACCESSION B1766766  
 VERSION B1766766.1 GI:15758344  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 REFERENCE 1 (bases 1 to 834)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE NIH-MGC National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/MLN at:  
 http://image.llnl.gov  
 Plate: L1AM1517 row: c column: 18  
 High quality sequence stop: 772.  
 Location/Qualifiers  
 1..834  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5206217"  
 /clone\_id="NIH\_MGC\_122"  
 /lab\_host="DH10B"  
 /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;  
 Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source

anonymous pool of 24 week female lung, 16 week female  
 spleen, and 20-22 week male spleens. Library is oligo-dt  
 primed and directionally cloned (ecory site is destroyed  
 upon cloning). Average insert size 1.4 kb, insert size  
 range 1-3 kb. Library is normalized and enriched for  
 full-length clones and was constructed by C. Gruber  
 (Invitrogen). Research Genetics tracking code 026. Note:  
 this is a NIH-MGC Library."

BASE COUNT 154 a 287 c 223 g 170 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1,19e-80 Length: 834  
 Score: 861.00 Matches: 172  
 Percent Similarity: 92.27% Conservative: 7  
 Best Local Similarity: 88.66% Mismatches: 14  
 Query Match: 74.10% Indels: 2  
 DB: 13 Gaps: 0

US-09-245-198A-2 (1-225) x B1766766 (1-834)

OY 33 GlnGluGluLeuThrAlaGluAspArgGluProProGlnLeuAsnProGlnThrGlu 52  
 DB 3 CAGAGAGAGTGTGGAGAGAGAGACAGACCCGTCGAGACTGAATCCACAGACAGA 62  
 OY 53 GluSerGlnAspValValProPheLeuGluGlnLeuValArgProArgArgSerAlaPro 72  
 DB 63 GAAGCCAGAGATCTCGCCCTTCTCTGAGACAGTTCAGTTCGCTGCGAGAGTGCACCT 122  
 OY 73 LysGlyArgGlyAlaArgProArgArgAlaIleAlaAlaHisTyrGlnValHisProArg 92  
 DB 123 AAAGCCGGAAGAAACAGGGCTCGAAGAGAGATCGCAGCCCATTAAGATTATCATCACA 182  
 OY 93 Pro-GlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTyrPglGluThr 112  
 DB 183 CCGTGACAGCAGCAGCAGCAGCAGCAGTGTGACGCGAGACAGTGTGCTGGAGAGAAC 242  
 OY 112 LysIleAsnSerSerSerProLeuArgTyrAspArginIleGlyGluPheThrValIle 132  
 DB 243 CAGAAATCAACAGCTCCAGCCCTGCTGCTACACGCGCAGATCGGGAGATTATAGTCA 302  
 OY 132 eArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyAlaValTyr 152  
 DB 303 CCGGCTGGGCTGTCTACTACCTGACTGAGTGCACCTTATGAGGGAGAGCTGTCTA 362  
 OY 152 IleuIleuAspLeuLeuValAsnGlyValLeuAlaLeuAlaGlyCysLeuGluPhe 172  
 DB 363 CCGTGAAGCTGGACTGT 422  
 OY 172 rAlaThrAlaAlaSerSerProGlyProGlnLeuArgLeuCysGlnValSerGlyLeu 192  
 DB 423 AGCCATCGGCGCAATTCCTCGGCGCCAGCTCCGCTTCCAGAGTGTCTGGACTGTT 482  
 OY 192 uProLeuArgProGlySerSerSerLeuArgIleArgThrLeuProTrrAlaHisLeuVal 212  
 DB 483 GGCCCTGGGGGCGAGGTTC-TCCCTGGAGATCGGACCCCTCGGCGCCATCTCAAGC 541  
 OY 212 aAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 225  
 DB 542 TGCCCTCTCCACACTACTGCGACTGCGACTCTTCAGGTTTCAC 581  
 RESULT 7  
 AM763237 561 bp mRNA linear EST 04-MAY-2000  
 LOCUS AM763237  
 DEFINITION ur70d09.y1 NCI\_CGAP\_Mam3 Mus musculus cDNA clone IMAGE:3155633 5'  
 similar to TR:054907 054907 TME-RELATED WEAK INDUCER OF APOPTOSIS  
 ;, mRNA sequence.  
 ACCESSION AM763237  
 VERSION AM763237.1 GI:7695174  
 KEYWORDS EST.  
 SOURCE house mouse.  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE  
1 (bases 1 to 561)  
AUTHORS  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
TITLE  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL  
Unpublished (1997)  
COMMENT  
Other\_ESTs: ur70d09.x1  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Lohar Hemighausen Ph.D., Chu-Xia Deng Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Washington University Genome Sequencing Center.  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
[image.llnl.gov/image/html/resources.shtml](http://image.llnl.gov/image/html/resources.shtml)

FEATURES  
source  
MGI:1058389  
Seq primer: -40RP from Gibco  
High quality sequence stop: 433.  
Location/Qualifiers  
1..561  
/organism="Mus musculus"  
/strain="129,C57BL/6J,FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:3155633"  
/clone\_id="NCI-CGAP\_Mam3"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="10 months"  
/lab\_host="DH108"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site\_1: Salt;  
Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT.  
Library constructed by Life Technologies. Investigators  
providing samples: Lohar Hemighausen/Chu-Xia Deng, NIH  
Reference for transgenic model: Xu et al., Nature Genetics  
22, 37-43 (1999)."

BASE COUNT 108 a 158 c 194 g 100 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1,77e-80 Length: 561  
Score: 857.00 Matches: 167  
Percent Similarity: 99.408 Conservative: 0  
Best Local Similarity: 99.408 Mismatches: 1  
Query Match: 73.75% Indels: 0  
DB: 10 Gaps: 0

US-09-245-198a-2 (1-225) x AM763237 (1-561)

QY 1 ValLeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeu 20  
|||||  
DB 58 GTGCTGAGCTGGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 117  
QY 21 GlySerTrpAlaThrLeuSerAlaGlnGlnProSerGlnGlnLeuThrAlaGlnASP 40  
|||||  
DB 118 GGGAGCTGGGCAAGCGCTGTGCCAGAGACCTTCTAGAGAGAGCGACAGCAGAGAC 177  
QY 41 ArgArgGluProProGluLeuAsnProGlnThrGlnGluSerGlnAsnValValProPhe 60  
|||||  
DB 178 CGCCGGAGAGCCCTGAGATCGATCCACACAGAGAGAGAGAGAGAGAGAGAG 237  
QY 61 LeuGlnGlnLeuValArgProArgArgSerAlaProLysGlnArgLysAlaArgProArg 80  
|||||  
DB 238 TTGGAGCAACTAGTCCGGCTCGAAGAGTGTCTTAAGGCGGAGAGGCGGCGCTCGC 297  
QY 81 ArgAlaLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 100  
|||||  
DB 298 CGAGCAATTCACAGCCCTTGTAGAGTTTCATCTCGGCGAGACAGAGTGAAGCACA 357  
QY 101 G1ValAspGlyThrValSerGlyTyrGlnGlnThrLysAlaAsnSerSerProLeu 120  
|||||  
DB 358 GGTGTGATGGACAGTGAAGTGGCTGGAGAGAGCAAAATCAACAGCTCCAGCCCTCTG 417

QY 121 ArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr 140  
|||||  
DB 418 CGGTACAGCCGACAGATTTGGGAATATACAGTCATGAGGGCTGGCTTACTACTGTAC 477  
QY 141 CysGlnValHisPheAspGlyGlyValValTyrLeuLysLeuAsnLeuValAsn 160  
|||||  
DB 478 TGTACAGTGCACCTTGTATGAGGAGAAAGCTGTACTTAAGCTGACCTGTGTGAC 537  
QY 161 G1ValLeuAlaLeuArgCysLeu 168  
|||||  
DB 538 GGTGTGCTGGCCCTGCGCTGCTG 561

RESULT 8  
LOCUS B0707185 948 bp mRNA linear EST 16-JUL-2002  
DEFINITION AGENCOURT\_8353983 NIH\_MGC\_113 Homo sapiens CDNA clone IMAGE:6278608  
5', mRNA sequence.  
ACCESSION B0707185  
VERSION B0707185.1 GI:21846084  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
1 (bases 1 to 948)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Dr. Mark Watson  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LICM2466 row: n column: 17  
High quality sequence start: 24  
High quality sequence stop: 550.  
Location/Qualifiers

FEATURES  
source  
1..948  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6278608"  
/clone\_id="NIH\_MGC\_113"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: spleen; Vector: pOTB7; Site\_1: XhoI; Site\_2:  
EcoRI; CDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGCAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using zap-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: This is a  
NIH\_MGC library."

BASE COUNT 173 a 334 c 252 g 187 t 2 others  
ORIGIN

Alignment Scores:  
Pred. No.: 1.9e-77 Length: 948  
Score: 831.50 Matches: 164  
Percent Similarity: 92.31% Conservative: 4  
Best Local Similarity: 90.11% Mismatches: 13  
Query Match: 71.56% Indels: 2  
DB: 14 Gaps: 1

US-09-245-198a-2 (1-225) x B0707185 (1-948)

QY 44 ProProGluLeuAsnProGlnThrGlnGluSerGlnAsnValValProPheLeuGln 63  
|||||  
DB 3 CCGTGGAACTGATCTCCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 62  
QY 64 LeuValArgProArgArgSerAlaProLysGlyArgLysAlaArgProArgArgAlaIle 83

Db	63	CTAGTTGGCGCTCGAAGAAAGTGCACCTTAAGCGCCGGAACACAGGGGCTCGAAGACCATC	122
Qy	84	AAAlAHs1YrGLuValHisProArqProGlyGlnAspGlyAlaGlnAlaGlyValAlaSp	103
Db	123	GCAGCCCATTTATGAAGTTCATCCACGACCTGGACAGGACGGAGC---GCAGGTGTGGAC	178
Qy	104	GLYHrVALsArGLYTrTGLUGluHrThrHisLeuAsnSerSerProLeuArqTrAsp	123
Db	179	GGGACAGTAACTGGCTGGAGGAGAACCCAGAAATCAACACCTCCAGCCCTCGCGCTAACAC	238
Qy	124	ArgGlnIleGlyGluPheThrValIleArqAlaGlyLeuTrpTyrTyrCysGlnAl	143
Db	239	CGCCAGATCGGGGAAATTATAGTCAACCGGGGTGGGCTCTACTACTACTACTAGGTG	298
Qy	144	HisPheAspGluGlyLysAlaValAlaTyrLeuLysLeuAspLeuValAsnGlyValLeu	163
Db	299	CACTTGATGAGGGGAAAGCGCTGTACTCCGTAAGCTGACTGTCTGGTGGATGGTGTGGC	358
Qy	164	AlaLeuArqCysLeuGluGluPheSerAlaHrAlaAlaAspSerProGlyProGlnLeu	183
Db	359	GGCTCGCGCTGTCTGAGAAATCTCACACATGGGCGCAAGTCTCCCTGGGCGCCAGCTC	418
Qy	184	ArgLeuCysGlnValSerGlyLeuLeuProLeuArqProGlySerSerLeuArqIleArq	203
Db	419	CGCCTCTCCAGAGTGTCTGGGCTGTGGCCCTGGCGGCGAGGGTCTCCCTGGAGATCCGC	478
Qy	204	ThrLeuProTrpAlaHisLeuLysAlaAlaAlaProPheLeuHrTyrPheGlyLeuPheGln	223
Db	479	ACCCCTCCCTGGGCGCAATCTCAAGGCTGCCCTCCCTCCACTACTGTGGACTCTTCAG	538
Qy	224	ValHis 225	
Db	539	GTTGAC 544	
RESULT 9			
BO884231			
LOCUS			
DEFINITION	BO884231	940 bp	mRNA linear EST 16-AUG-2002
ACCESSION	AGNC00097	8682031	Lupsk1.sc1atic.nerve Homo sapiens .cDNA clone
VERSION	IMAGE:6197488	5'	mRNA sequence.
KEYWORDS	BO884231		
SOURCE	BO884231.1	GI:22276239	EST.
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.		
JOURNAL	1 (bases 1 to 940)		
COMMENT	NIH-MGC htcp://mgc.nci.nih.gov/.		
	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgaabs-remail.nih.gov		
	Tissue Procurement: Dr. James R. Lupski		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	htp://image.lnl.gov		
	plate: LLM13607 row: 1 column: 17		
	High quality sequence stop: 453.		
FEATURES	Location/Qualifiers		
source	1..940		

```
1..940
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="6197488"
/clone_id="lupski_sciatic_nerve"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DR10B"
note=Vector: pCMV-Sport6 (Life Technologies); Site 1
```

North: Site 2: Salt; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCACGCGTCG-3' and 5'-GACTAGTTCTAGATCGGAGAGCGGCCCG(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies.<sup>a</sup>

BASE COUNT	170 a	325 c	255 g	190 t
------------	-------	-------	-------	-------

Alignment Scores:	
Pred. No.:	5, 03e-76
Score:	818.00
Percent Similarity:	92.02%
Best Local Similarity:	88.30%
Query Match:	70.40%
DB:	14
Length:	940
Matches:	166
Conservative:	7
Mismatches:	11
Indels:	0
Gaps:	0

US-09-245-198A-2 (1-225) x BQ884231 (1-940)

Oy		42	ArgGluProProGlnIleuAsnProGlnThrGluGluSerGlnAspValValProPheLeu	61
Db		1	CAGGACCCGGTGGAACTGAATCCCGACAGAGAAGAACCCAGGATCCTGGCTTTCTGTG	60
Oy	GluGlnIleuValArgProArgArgSerAlaProLysGlyArgLysAlaArgProArgArg	81		
Db	AACCGACTAGTTGGGCCCTGCAGAAAGTCACCTAAGGCCCGGAAAAACAGGGGCTGGAAGA	120		
Oy	AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly	101		
Db	GGCATCCGACGCCATTAAAGAGTTTCATCCACAGACCCTGACAGGAGCGAGCGACGACGT	180		
Oy	ValAspGlyThrValSerGlyTyrPgluglunThrLysIleAsnSerSerProLeuArg	121		
Db	GTGGACGGGACAGTAGTGAGTGGCTGGAGGAGACCAAGATCAACAGCTCCAGCCCTTGGCC	240		
Oy	TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys	141		
Db	TACAACCGCCAGATCGGGAGGAGTTTATGATCACCCGGGCTGGGCTTACTACTGTACTGT	300		
Oy	GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspPleuLeuValAsnGly	161		
Db	CAGGCGCACTTGTGATGAGGGGAGAGGCTGTCTACCTGAAGCGTGGACTTCTGTGTGATGT	360		
Oy	ValIleuAlaIleuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyPro	181		
Db	GTGTGGGGCCCTGGCGCTGCTGTGGAGGAATTCACGCACTGGCGGCAAGTTCCCTCGGCCC	420		
Oy	GlnIleuArgLeuCys-GlnValSerGlyLeuLeuProLeuArgProGlySer-SerLeuA	201		
Db	CAGCTCCGCCCTCTGGCCACAGGTGTCTGGGCTGTGGCCCTGGGGCCAGGGGCCCTCCCTGC	480		
Oy	TgIleArgThrLeuProThrAlaHisLeuLysAlaAla-ProPheLeuThrTyr-PheGl	220		
Db	GGATCCGACACCTCCCTCGGGGCCATCTCAAGAGCTGCCCCCTTCTACCTACTTTCGG	540		
Oy	YleuPheGlnValHis	225		
Db	ACTCTTCACGGTTCAC	556		
RESULT 10				
LOCUS	B0671259			
DEFINITION	B0671259	963 bp	mRNA	linear EST 15-JUL-2002
ACCESSION	AGNCOCURT_8303564	NIH_MGC_102	Homo sapiens cDNA clone IMAGE:6274716	
VERSION	B0671259	.1	GI:21782093	
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

REFERENCE 1 (bases 1 to 963)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: ggapbs-remail.nih.gov  
 Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LLCM2456 row: 1 column: 13  
 High quality sequence stop: 565.  
 Location/Qualifiers

## FEATURES

## SOURCE

1. 963  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_1db="NIH\_MGC\_102"  
 /issue\_type="epidermoid carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: salivary gland; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; CDNA made by oligo-dt priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

## BASE COUNT

194 a 326 c 260 g 183 t

## ALIGNMENT SCORES:

Pred. No.: 2.69e-70 Length: 963  
 Score: 764.00 Matches: 151  
 Percent Similarity: 95.65% Conservative: 3  
 Best Local Similarity: 93.79% Mismatches: 7  
 Query Match: 65.75% Indels: 1  
 DB: 14 Gaps: 0

US-09-245-198a-2 (1-225) x BQ671259 (1-963)

65 VALARGPARGARGSERALAPROLYSGLYARGLYSALAARGPROARGARGALAILEALA 84  
 |||||||  
 2 GTTCGGCCTCGCAGAAAGTGCACCTAAAGCGGAAAAA-CGGGCTCGAAGAGCGATCGCA 60  
 |||||||  
 85 AAlAlstYrGlUvAlHsPrgArGPrGgLyGlnAsPgLyAlAGlnAlAGlyVAlAsPgLy 104  
 |||||||  
 61 GCCCATTATGAAGTTCAATCCACGACCTGGACAGCAGCAGCGAGCGAGCGGTGTGACGG 120  
 |||||||  
 105 ThrVAlSerGlyTrPGlUGlUThrTlSleAsnSerSerSerProLeuArgTrAspArg 124  
 |||||||  
 121 ACACTGAGTGGCTGGGAGAAAGCCAGAAATCAACAGCTCCACCTCTGGCGTTACAAACGGC 180  
 |||||||  
 125 GlnlleglGluPheThrVAlleArgAlAGlyLeuTrYTrYleuTrYCyGlnVAlHs 144  
 |||||||  
 181 CAGATCGGGGAGTTATATAGTACCCGGGCTGGCTACTACTACTGTCAGATGTCAGTGCAC 240  
 |||||||  
 145 PheAsPgLyGlyLysAlaVAlTrYleuLysLeuAspLeuLeuVAlAsnGlyVAlleuAla 164  
 |||||||  
 241 TTTGATAGGGGAGGAGGCTGTCTACTGTAAAGCTGACCTGCGTGGATGGTGTCTGCGCC 300  
 |||||||  
 165 LeuArgCySLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyProGlyProGlyLeuArg 184  
 |||||||  
 301 CTGGCGTGGCTGGAGAAATTTCAACCACTGGCGCAATTCCTCTGGGCGCCACCTCCGCG 360  
 |||||||  
 185 LeuCySglnVAlSerGlyLeuLeuProLeuArgProGlySerSerLeuArgTrIleArgThr 204  
 |||||||  
 361 CTCGCGCAGAGTGTCTGGGCTGTGGCCCTGGCGCAGAGGTCTCTCCCTCGGAGATCGCACCC 420  
 |||||||

QY 205 LeuProTrpAlaHsLeuYsAlaAlaProPheLeuThrTrpPheGlyLeuPheGlnVAl 224  
 |||||||  
 Db 421 CTCCTCGTGGCCATTCACAGGCTGGCCCTTCTCTCACCTACTTCGGACCTCTTCACAGTT 480  
 |||||||  
 QY 225 Hts 225  
 |||  
 Db 481 CAC 483

RESULT 11  
 AK020909 1033 bp mRNA linear HTC 19-JAN-2002  
 LOCUS Mus musculus adult retina cDNA, RIKEN full-length enriched library,  
 DEFINITION clone:A930030D13:tumor necrosis factor (ligand) superfamily, member  
 12, full insert sequence.

ACCESSION AK020909  
 VERSION AK020909.1 GI:12861640  
 KEYWORDS HTC; CAP trapper.  
 SOURCE Mus musculus (strain: C57BL/6J) adult retina cDNA to mRNA,  
 clone\_1db:RIKEN full-length enriched mouse cDNA library  
 clone:A930030D13.  
 Mus musculus

## ORGANISM

REFERENCE 1  
 AUTHORS Carninci, P. and Hayashizaki, Y.  
 TITLE High-efficiency full-length cDNA cloning  
 JOURNAL Meth. Enzymol. 303, 19-44 (1999)  
 MEDLINE 99279253  
 PUBMED 10349636

REFERENCE 2  
 AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.  
 TITLE Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes  
 Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL MEDLINE 20499374  
 PUBMED 11042159  
 REFERENCE 3  
 AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Sasaki, N., Carninci, P.,  
 Kono, H., Akiyama, J., Nishii, K., Kitsuana, T., Tashiro, H., Itoh, M.,  
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,  
 Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohtsuka, E., Watanabe, M.,  
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuda, S., Kawai, J.,  
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.  
 TITLE RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)

JOURNAL MEDLINE 20530913  
 PUBMED 11076861  
 REFERENCE 4  
 AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,  
 Arakawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S.,  
 Aizawa, K., Izawa, M., Nishii, K., Kiyosawa, H., Kondo, S., Yamane, I.,  
 Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,  
 Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,  
 Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,  
 Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,  
 Quackenbush, J., Schriml, L.M., Stauble, F., Suzuki, R., Tomita, M.,  
 Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,  
 Baldarelli, R., Barsby, G., Blake, J., Boffelli, D., Bojunga, N.,  
 Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bull, C.,  
 Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,  
 Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,  
 Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P.,  
 Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,  
 Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,  
 Toyokawa, K., Wang, K.H., Weitz, C., Whitaker, C., Wilming, L.,  
 Wyshaw-Borls, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S.,  
 and Hayashizaki, Y.  
 TITLE Functional annotation of a full-length mouse cDNA collection  
 Nature 409 (6821), 685-690 (2001)  
 JOURNAL MEDLINE 21085660

PUBMED  
 REFERENCE  
 11217851  
 5 (bases 1 to 1033)  
 Authors  
 Aadachi, J., Alizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,  
 Arikawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C.,  
 Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanganbuchi, T.,  
 Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hiraoka, T., Hori, F.,  
 Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kaakawa, T.,  
 Kato, H., Kawaji, J., Kojima, Y., Konno, H., Kouda, M., Koya, S.,  
 Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K.,  
 Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J.,  
 Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,  
 Schriml, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T.,  
 Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F.,  
 Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I.,  
 Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M., and  
 Hayashizaki, Y.  
 Title  
 Direct Submission  
 Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of  
 Physical and Chemical Research (RIKEN), Laboratory for Genome  
 Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
 RIKEN Yokohama Institute, 1-7-22 Saito-cho, Tsukuba-shi, Yokohama,  
 Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp,  
 URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,  
 Fax: 81-45-503-9216)  
 Comment  
 Please visit our web site (http://genome.gsc.riken.go.jp/) for

```

FEATURES
source
location/Qualifiers
1..1033
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:A930030D13"
/db_xref="MGI:MGI:1911867"
/db_xref="taxon:10090"
/clone="A930030D13"
/issue_type="retina"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1..1033
/gene="Tnfrsf12"
<1..433
/gene="Tnfrsf12"
/note="data source:MGI, source key:MGI:1196259,
evidence:ISS
putative
tumor necrosis factor (ligand) superfamily, member 12"
/codon_start=2
/protein_id="BAB32249.1"
/db_xref="GI:12861641"
/db_xref="MGI:MGI:1196259"
/translation="IAAHYEVHPRPGDGAQAGVDCGYSSGEETKINSSSLRDROQ
GEFVIVAGLILYLYCQVHFDEGKAVYILKLDLIVGVLAALRCLFEESATPAASSPGQLR
LCQVSGILPLNPGSSLRIRTLPMWHLAAAPLVEYFGLFQVH"

```

BASE COUNT	228 a	300 c	248 g	256 t	1 others
ORIGIN					
Alignment Scores:					
Pred. No.:	3,4e-69	Length:	1033		
Score:	754.00	Matches:	143		
Percent Similarity:	100.00%	Mismatches:	0		
Best Local Similarity:	100.00%	Indels:	0		
Query Match:	64.89%	Gaps:	0		
US-09-245-198A-2 (1-225) x AK020909 (1-1033)					
QY	11eAlaAahistYrGuValaHisProArGProGlyGlnAspGlyAlaGlnAlaGlyVal	102			
Db	2 ATTCACACCCCATTTGAGGTTTCATCTCGGCCAGACAGATGAGACACAAGAGGTGTG	61			
QY	103 AspelGhrValSerGlyTTPGluGluThrTrpHisIleAsnSerSerProLeuArGlyTr	122			
Db	62 GATGGACAGTGAAGTGGCTGGGAAGACCAAAATCAACACCTCCAGCCCTTCGGCTAC	121			
QY	123 AspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTrpTyrTyrCysGln	142			
Db	122 GACCGCCAGANTGGGGAATTACAGTCATCAGGGCTGGGCTCTACTACTACTACTAG	181			
QY	143 ValHisPheAspGluGlyLysAlaValIleuLysLeuAspLeuValaAspGlyVal	162			
Db	182 GTGACACTTGTATGAGGGAAGAGGCTGTCTCACTGAAGCTGAGCTGGTGAAGGTGTG	241			
QY	163 LeuAlaLeuArGySLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyProGln	182			
Db	242 CTGGCCCTGGCTCTCTCTGGAGAAATTTCTACGCCACAGCAGAAAGCTCTCTGGGCCGAG	301			
QY	183 LeuArGLeuCySglnValSerGlyLeuLeuProLeuArGProGlySerSerLeuArgIle	202			
Db	302 CTCGGTTTGCCAGGAGTGTCTGGGCTGTGGCCGTGGCGGACAGGATCTTCCCTTCGATTC	361			
QY	203 ArgThrLeuProTrrPalHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPhe	222			
Db	362 CGCACCCCTCCCTGGGCTCACTTAAAGCTTCGCCCTTCCTCAACCTACTTTGGACTCTTT	421			
QY	223 GlnValHis 225				
Db	422 CAACTTCAC 430				
RESULT 12					
LOCUS	BM921213	1071 bp	mRNA	linear	EST 12-MAR-2002
DEFINITION	AGENCOURT_6633046 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752561				
ACCESSION	BM921213				
VERSION	BM921213.1				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
TITLE	1 (bases 1 to 1071)				
JOURNAL	NIH-MGC http://imgc.ncl.nih.gov/				
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Unpublished (1999)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgaabs-remail.nih.gov				
	Tissue Procurement: Life Technologies, Inc.				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)				
	Clone Sequencing by: Agencourt Bioscience Corporation				
	DNA distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LNL at:				
	http://image.lnl.gov				
	plate: ILM12786 row: p column: 02				
	High quality sequence stop: 656.				
	Location/Qualifiers				



```

Db      422  CTTCCCGGGGGCCATTAAAGGAGTGCCCTTCTCCTCACCATTGTTGAATTTT 477
RESULT 14
LOCUS    AM917574                      584 bp    mRNA    linear    EST 25-MAY-2000
DEFINITION EST346878 Rat gene index, normalized rat, norvegicus, Bento Soares
Rattus norvegicus cDNA clone RGIIEF4 5' end, mRNA sequence.
ACCESSION AM917574.1 GI:8083328
VERSION   AM917574.1
KEYWORDS  EST.
SOURCE    Norway rat.
ORGANISM  Rattus norvegicus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
           Rattus.
REFERENCE 1 (bases 1 to 584)
AUTHORS   Lee,N.H., Glodet,A., Chandra,I., Mason,T.M., Quackenbush,J.,
           Kerlavage,A.R. and Adams,M.D.
           Rat Genome Project: Generation of a Rat EST (RESE) Catalog & Rat
           Gene Index
TITLE      Unpublished (1998)
JOURNAL    Contact: Lee, NH
COMMENT    The Institute for Genomic Research
           9712, Medical Center Drive, Rockville, MD 20850, USA
           Tel: (301)-838-3529
           Fax: (301)-838-0208
           Email: nhlee@tigr.org
           This clone is available through the ATCC, contact the ATCC
           tel#703-365-2700 for further information
           Seq primer: M13 Reverse.
FEATURES
  source   1..584
            /organism="Rattus norvegicus"
            /db_xref="taxon:10116"
            /clone="RGIIEF4"
            /clone_id="Rat gene index, normalized rat, norvegicus,
            Bento Soares"
            /tissue-type="mix - brain, ovary, placenta, kidney, lung,
            liver, embryo, heart, muscle, spleen"
            /lab_host="SOLR"
            /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
            XhoI; Estimated insert size approx.1 kb"
BASE COUNT 108 a 188 c 147 g 138 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 9.72e-64 Length: 584
Score: 699.00 Matches: 133
Percent Similarity: 99.26% Conservative: 1
Best Local Similarity: 98.52% Mismatches: 1
Query Match: 60.15% Indels: 0
Gaps: 0
US-09-245-198a-2 (1-225) x AM917574 (1-584)
Oy      91  ProaagpPpoglyGlnaspglyAlaGlnaGlyValaspglyThrValserGlyTrrpGlu 110
Db      2  CCACGCCAGGACAGGATGAGACACAGGAGGTGATGGAGGAGTGGCTGGGAA 61
Oy      111 GluthrlyslleasnserserserProleuargtyrAsparGlnleGlyslunpethr 130
Db      62  GAGACCAAAATCAACAGCTCCAGCCCTCGCCCTATGACCGCCACATTTGGGAATTTAG 121
Oy      131 VallearGalaGlyLeuTyrrTyrrleuTyrrCysGlnValHisPhaspsGlnGlyTysala 150
Db      122 GTCAACAGGCGCTGAGCTTACTACTGCTGACGTGACGTGACGTGATGAGGAGGAGCA 181
Oy      151 ValtyrleuTyrsleuasleuValasnglyValleuAlaalaLeuargCysleuGlnu 170
Db      182 GCTACCTGAAGAGCTGAGCTGCTGATGATGCTGGCCCTCGCTGGAGAAGA 241
Oy      171 PheseralarhrAlaAlaserSerProGlyProGlnleuArgleuGlyGlnValserGly 190

```

```

Db      242  TTCTCAGCCACGACGAGCGAGCTCTCGGGGCCCGCCAGCTCCGCTGCGAGGTGCTGGG 301
Oy      191  LeuleupProleuArgProGlySerSerleuArgllleargThleuProTrrAlaHisleu 210
Db      302  CTGTGGCTCTGTCGGCCGAGGAGGTCTTCCCTTCGATCCGATCCCTCCCGGGGCTCATCTT 361
Oy      211  LysAlaAlaProPhelenuThrrPhrheGlyLeuPhneGlnValHis 225
Db      362  AAGGCTGCCCTCTCTTACTTGTGAGACTTNTCAAGTTTAC 406
RESULT 15
LOCUS    AA870722                      445 bp    mRNA    linear    EST 16-MAR-1998
DEFINITION vq25g07.21 Barstead stromal cell line MRLR8 Mus musculus cDNA
LOCUS IMAGE:1095324 5', mRNA sequence.
ACCESSION AA870722
VERSION   AA870722.1 GI:2966167
KEYWORDS  EST.
SOURCE    house mouse.
ORGANISM  Mus musculus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 445)
AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
           Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
           Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
           Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
           Waterston,R.
           The WashU-HHMI Mouse EST Project
TITLE      Unpublished (1996)
JOURNAL    Contact: Marra M/Mouse EST Project
COMMENT    WashU-HHMI Mouse EST Project
           Washington University School of MedicineP
           4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
           Tel: 314 286 1800
           Fax: 314 286 1810
           Email: mouseest@watson.wustl.edu
           This clone is available royalty-free through LLNL; contact the
           IMAGE Consortium (info@image.llnl.gov) for further information.
           Seq primer: -28m13 rev2 ET from Amersham
           High quality sequence stop: 420.
FEATURES
  source   1..445
            /organism="Mus musculus"
            /db_xref="taxon:10090"
            /clone="IMAGE:1095324"
            /clone_id="Barstead stromal cell line MRLR8"
            /cell_line="C2C12 (undifferentiated)"
            /lab_host="DH10B"
            /note="Vector: p7T730-Pac (Pharmacia) with a modified
            polylinker; Site_1: EcoRI; Site_2: NotI; 1st strand cDNA
            was primed with a Not I - oligo(dT) primer [5'
            TGTACGACATCTGAAGTGGAGCGCGCCCGCTTTTCTTTTCTTTTCTTTTCTTTT
            3']; double-stranded cDNA was ligated to Eco RI adaptors
            [AATTCGATCTCTG], digested with Not I and cloned into the
            Not I and Eco RI sites of the modified p7T73 vector.
            Source undifferentiated tissue culture cell line C2C12.
            Library constructed by Bob Barstead. The C2C12 cell line
            (available from ATCC, catalog # CRJ-1772) differentiates
            rapidly, forming contractile myotubes and producing
            characteristic muscle proteins.
BASE COUNT 102 a 119 c 135 g 89 t
ORIGIN
Alignment Scores:
Pred. No.: 7.64e-61 Length: 445
Score: 670.00 Matches: 132
Percent Similarity: 96.38% Conservative: 1
Best Local Similarity: 95.65% Mismatches: 5
Query Match: 57.66% Indels: 1
Gaps: 0

```

us-09-245-198a-2 (1-225) x AA870722 (1-445)

```
OY 50 GlnThrGluGluSerGlnAspValProPheLeuGluGlnLeuValArgProArg 69
    |||
    |||
    |||
Db 12 CAAGCGTCGACAGCGCAGATGTGTACTTCTTGAAACACTAGTCCGGCTCGAAGA 71
    |||
OY 70 SerAlaProLysGlyArgLysAlaArgProArgArgAlaIleAlaAlaHisTyrGluVal 89
    |||
    |||
    |||
Db 72 AGTGCTCCTAAAGGCCGGAACGG-CGGCCTCGCCGAGCTATTGCAGCCCATATGAGGTT 130
    |||
OY 90 HisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTyr 109
    |||
    |||
    |||
Db 131 CATCCTCGGCCGACGACAGATGACACACAGCAGGTGTGATGGACACAGTGTGCTGG 190
    |||
OY 110 GluGluThrLysIleAsnSerSerSerProLeuArgTyrAspArgGlnIleGlyLuphe 129
    |||
    |||
    |||
Db 191 GAAGAGACCAAAATCAACAGCTCCAGCCCTCGCCTACGACCGCCAGATTGGGAATTT 250
    |||
OY 130 ThrValIleArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyLys 149
    |||
    |||
    |||
Db 251 ACAGTCATCAGGGCTGGGCTCTACTACTGTAAGTGTGACAGGTGCACTTGATGAGGAAAG 310
    |||
OY 150 AlaValTyrLeuLysLeuAspLeuValAsnGlyValIleuAlaLeuArgCysLeuGlu 169
    |||
    |||
    |||
Db 311 GCTGCTACCTGAAGCTGACTGTGCTGTAACGGGTGTGCTGCGCTGCGCTGCGGAA 370
    |||
OY 170 GluPheSerAlaThrAlaAlaSerSerProGlyProGlnLeuArgLeuCysGln 187
    |||
    |||
    |||
Db 371 GAATTCCTCAGCCACAGCAGCAAGCTCTCTGGGCCCGCCAGCTCCGTTGTGCCAG 424
```

Search completed: May 8, 2003, 04:22:48  
Job time : 1064.98 secs



GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:11:52 ; Search time 23.8703 seconds

(without alignments)  
1942.185 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162  
Sequence: 1 VLSGLALACGLLVVSL.....PMAHLKAPFLTYEGLFGVH 225

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mhc:\*  
8: sp-organelle:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*  
15: sp-virus:\*  
16: sp-bacteriap:\*  
17: sp-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107.5	9.3	325	5	09V5G2 drosophila
2	93.5	8.0	557	16	08XGX3 raietonia s
3	90.5	7.8	210	16	09A9Z6 caulobacter
4	90.5	7.8	421	16	09H0W2 pseudomonas
5	89.5	7.7	805	5	09VFD4 drosophila
6	89	7.7	287	13	090WT9 gallus galli
7	89	7.7	352	12	089341 hendra viru
8	88	7.6	224	5	09Y762 drosophila
9	88	7.6	353	12	066760 equine morb
10	88	7.6	532	4	016727 homo sapien
11	88	7.6	1663	4	09U001 homo sapien
12	88	7.6	3261	4	09Y556 homo sapien
13	88	7.6	3664	4	096T58 homo sapien
14	87.5	7.5	522	10	09FTV7 oryza sativ
15	87.5	7.5	670	16	09AAL5 caulobacter
16	86.5	7.4	340	16	09H0R8 pseudomonas

17	86.5	7.4	504	16	092KA4 rhizobium m
18	86	7.4	260	10	08S2N9 oryza sativ
19	85.5	7.4	331	10	0942P9 oryza sativ
20	84.5	7.3	422	16	09RK50 streptomyce
21	84.5	7.3	430	2	09REU1 streptomyce
22	84.5	7.3	1079	13	08UVR4 xenopus lae
23	84.5	7.3	1118	16	09BE34 rhizobium l
24	84	7.2	2962	5	093326 caenorhabd
25	83.5	7.2	394	16	092V66 rhizobium m
26	83	7.1	467	16	09S2T4 streptomyce
27	83	7.1	2779	5	09W4N7 drosophila
28	82.5	7.1	174	16	09CKX1 pasteurella
29	82.5	7.1	314	5	08WPH7 theileria a
30	82.5	7.1	549	16	08RC38 thermococ
31	82	7.1	619	5	08S0H9 encephalito
32	82	7.1	654	16	P74654 synecocyst
33	81.5	7.0	999	11	09JXR6 mus musculu
34	81.5	7.0	1665	11	08V111 mus musculu
35	81	7.0	258	4	09NUDO mus musculu
36	81	7.0	262	11	09CVL1 mus musculu
37	80.5	6.9	274	10	0949I6 beta vulgar
38	80.5	6.9	324	11	09JUS4 rattus norv
39	80.5	6.9	788	4	09P2E5 homo sapien
40	80.5	6.9	1438	16	085739 pseudomonas
41	80	6.9	319	16	08U8B6 agrobacteri
42	80	6.9	403	6	09BGL8 canis famli
43	80	6.9	589	5	08T0B2 drosophila
44	80	6.9	793	5	09Y108 drosophila
45	80	6.9	1522	10	08S4C3 oryza sativ

## ALIGNMENTS

RESULT 1	ID	09V5G2	PRELIMINARY:	PRT:	325 AA.
AC	09V5G2	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)				
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)				
DE	CG12919 protein.				
GN	CG12919.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC	Ephydroidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BERKELEY:				
RX	MEDLINE=20196006; PubMed=10731132;				
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,				
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,				
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champs M., Pfeiffer B.D.,				
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,				
RA	Abdill J.F., Agdayani A., An H.-J., Andrews-Pfannkoch C., Batlowin D.,				
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,				
RA	Bortova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,				
RA	Burtis K.C., Busam D.A., Butler H., Cadieux E., Center A., Chandra I.,				
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,				
RA	Durkin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,				
RA	Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,				
RA	Glodek A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,				
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,				
RA	Hastin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwan C.,				
RA	Jajali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum R.A.,				
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,				

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Relbert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shier B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spletter E., Spredding A.C., Stapleton M., Strong R., Sun E.,  
 RA Svetskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003831; AAF58848.1;  
 DR FlyBase: FBgn003483; CG12919.  
 DR InterPro: IPR000478; TNE\_family.  
 DR SMART: SM00207; TNE\_1;  
 DR PROSITE: PS00251; TNE\_1;  
 DR PROSITE: PS00459; TNE\_2;  
 SQ SEQUENCE 325 AA; 36862 MW; 6E5CBB69694F1A3A CRC64;

Query Match 9.3%; Score 107.5; DB 5; Length 325;  
 Best Local Similarity 23.1%; Pred. No. 0.046; Indels 35; Gaps 9;  
 Matches 48; Conservative 34; Mismatches 91;

QY 29 QEPSEELTAEDRRPELPQTESQDVVPLEQLVPRRS--APKGRAPRAIAAH 86  
 DB 141 QEKSSNEATSKESPAPLHHRMRHNRH-----RHLLVRGSELARSDSRP----AAH 191  
 QY 87 YEYHRRPGDQAGAVDQVSGMEETKINSSPLRYDQIGFTYIRAGLYLYLCQVHPD 146  
 DB 192 FHLSRRRRHOGSM-GYHDMYIGNDNERNISYOG-HFQTRDGYLTATNGLLYYVYQICYN 249  
 QY 147 EGRKAVYLTLDLVNGVLA-----LRCLFEFSATPAASPQGLTLCQVSGLLPRPSS 199  
 DB 250 NSHD-----QNGFIYFGDTPPLQCLN-----TVPTNMPKAVTCHTSGLIHERNER 297  
 QY 200 LRIPLT---PWAHLKADPFLYFGLFQV 224  
 DB 298 IHLKDINHNRNAVLEGGNNRSYFGIFKY 325

## RESULT 2

ID 08XOX3 PRELIMINARY; PRT; 557 AA.  
 AC 08XOX3;  
 DT 01-MAR-2002 (Tremblrel. 20, Created)  
 DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Probable activation/secretion signal peptide protein.  
 GN RSP1095 OR NS02601.  
 OS *Ralstonia solanacearum* (Pseudomonas solanacearum).  
 OG *Plasmid megaplasmid*.  
 OC Bacteria; Proteobacteria; beta subdivision; *Ralstonia* group;  
 CC *Ralstonia*.  
 OX NCBI\_TaxID=305;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GM1100;  
 RX MEDLINE-21681879; PubMed-11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Chailat M., Billault A., Brottier P., Camus J.C., Catolico L.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
 RA Sigulier P., Thebaud P., Whalen M., Wincker P., Levy M.,  
 RA Weissbach J., Boucher C.A.,  
 RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*.";  
 RL Nature 415:497-502(2002).  
 DR EMBL: AL646082; CAD18246.1;

KW Plasmid; Complete proteome.  
 SQ SEQUENCE 557 AA; 60786 MW; 8B974C147D710649 CRC64;

Query Match 8.0%; Score 93.5; DB 16; Length 557;  
 Best Local Similarity 26.6%; Pred. No. 2;  
 Matches 62; Conservative 25; Mismatches 77; Indels 69; Gaps 14;

QY 2 LSLGLACGLLLVVSIGSMATLS-AQPSQELTAEDRRPE--LNPQTESQDVV 58  
 DB 10 LAGVALA-----GPLSSWAQSSPAGNPDIDLPVDTSRPEQKHVQGRPN--- 58  
 QY 59 PLEQL-----VPRRSAPGRKARRRAIAAHYEVHRRPGDQAGVDTVSGMEETKI 114  
 DB 59 PALENLATHLPSPKRFQIEGVATLPEPEIAHFA--PMAGHDVTAQLQAAAS--EYTKL 114  
 QY 115 -NSSPLRYDQIGFTYIRAGLYLYLCQVHFDGKAVYLTLDLVNGVLAQLCLEERS 172  
 DB 115 YADRGVPL-----SFAVVPQTF-----EGVVR--TVESYVARMRE--- 152  
 QY 173 ATAASSPGP---QLRLC-----QVSGLLPRPSSSLRIPLTP 206  
 DB 153 ----GKPGPLEARLNAISKHMDERPLRRETFERTVGLVLAQPGVQITATVP 201

## RESULT 3

ID 09A926 PRELIMINARY; PRT; 210 AA.  
 AC 09A926;  
 DT 01-JUN-2001 (Tremblrel. 17, Created)  
 DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Hypothetical protein CC1168.  
 GN CC1168.  
 OS *Caulobacter crescentus*.  
 OC Bacteria; Proteobacteria; alpha subdivision; *Caulobacter* group;  
 CC *Caulobacter*.  
 OX NCBI\_TaxID=155892;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19089 / CB15;  
 RX MEDLINE-21173698; PubMed-11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Ullendack T., Tran K., Wolf A., Yamathavan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,  
 RT "Complete genome sequence of *Caulobacter crescentus*.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005794; AAK23152.1;  
 DR TIGR: CC1168;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 210 AA; 22344 MW; 8B9830ADFBF7F45C CRC64;

Query Match 7.8%; Score 90.5; DB 16; Length 210;  
 Best Local Similarity 26.2%; Pred. No. 1.1;  
 Matches 33; Conservative 19; Mismatches 67; Indels 7; Gaps 2;

QY 2 LSLGLACGLLLVVSIGSMATLSAQPSQELTAEDRRPELPQTESQDVVPL 61  
 DB 32 LRLGLAVFAGAVATVYQAAMRSKLQTOQAATAPLVLDKPRF--TGVLKDGRRPL 88  
 QY 62 EQLVPRRSAPGRKARRRAIAAHYEVHRRPGDQAGV---DGTVSGMEETKINSS 117  
 DB 89 ITARARBDKADQIVLTVAPLVRYGGEPRPSGATKSGVYRAENTLLTLDEVKITSA 148  
 QY 118 SPLRYD 123  
 DB 149 EGDFD 154

## RESULT 4

```

09HUM2
ID 09HUM2 PRELIMINARY: PRT: 421 AA.
AC 09HUM2:
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Hypothetical protein PA4851.
GN PA4851.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision: Pseudomonadaceae;
OC Pseudomonas.
NCBI_TaxID=287;
RX STRAIN-ATCC 15692 / PA01:
RC MEDLINE-20437337; PubMed-10984043;
RA Stover C.K., Phan X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hutnagle W.O., Kowalik D.J., Lagrou M.,
RA Gardner R.L., Goltzy L., Tolentino E., Westbrook-Wedman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lam R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL; AE004898; AAC08236.1;
KV Hypothetical protein; Complete proteome.
SQ SEQUENCE 421 AA; 46442 MW; 68C5C17099953C3E CRC64;

Query Match 7.8%; Score 90.5; DB 16; Length 421;
Best Local Similarity 22.6%; Pred. No. 2.7;
Matches 53; Conservative 27; Mismatches 76; Indels 79; Gaps 10;

OY 29 QPSPQELTAEDERR-----REPELNPQTEESQDVPEL-----EQLV 65
DB 163 QPSPQELTAEDERR-----REPELNPQTEESQDVPEL-----EQLV 65
OY 66 RRRSAPGRKARPRRAIAHYEVHPRGQDGAQAGVDTGVSWEETKINSSSPRLYDRQ 125
DB 223 EQGSAVPLGSARKPDEA-----PVESDQLRDEPLQJLAWKEPR-----RQMPRR 266
OY 126 IGEFTVY---RAGLYTYCQVHDE-----GKAVYLKLDL--LVNGV 162
DB 267 LGMLLLILALGCLAAQYTAHYHDELARQDAPWFAQLCPETIGCLPSKYDVEQLRSSN 326
OY 163 LALRCLLEESA-----TAASSPGP--QLRLCYSGLL-----PLRPSSL 200
DB 327 LVVRSHPEFSGALVVDALITYNRASFSQPPLELRLRADLNGHLIANRRKPEYTL 381

RESULT 5
O9VFD4 PRELIMINARY: PRT: 805 AA.
AC 09VFD4:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE CG18442 protein.
GN CG18442.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Phlebotomina; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RX MEDLINE-20196006; PubMed-10731132;
RC STRAIN-BERKELEY;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner K., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

```

```

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.C.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.D., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibeigwan C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
RA Palazzo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL EMBL; AE003707; AAF55126.1;
DR Flybase; FBgn0038287; CG18442.
SQ SEQUENCE 805 AA; 86321 MW; FDF9EEFE14E69957 CRC64;

Query Match 7.7%; Score 89.5; DB 5; Length 805;
Best Local Similarity 21.7%; Pred. No. 7.7;
Matches 49; Conservative 27; Mismatches 55; Indels 95; Gaps 10;

OY 29 QPSPQELTAEDERR-----REPELNPQTEESQDVPELQVRRSAPGRKARPRR--ATAAH 86
DB 482 EQGSAVPLGSARKPDEA-----PVESDQLRDEPLQJLAWKEPR-----RQMPRR 266
OY 87 YEYHR-----FGDGAQAGVDTGVSWEETKINSSSPRLYDRQ 115
DB 530 PYMQPQERRELCPOLPPRGSPPLDGSOSPTNAVSGPKPPLPIACRPRPSNGVNSPN 589
OY 116 SSS-----PLRYDROI-----GEFTVIRAGLYTYCQV 143
DB 590 SSSPSGAPPRAHSPPIPAIVRPLHLNQAANGTLPPLPKKQOLHGEGLFTKNG----- 641
OY 144 HFDEGKAVYLKLDLVNGVL-----ALRCLLEFSATAASSPGPOL 183
DB 642 HIDEGLALLAKTDVAMSGILLIKLDVAAQC-----SVAQAAGGCTSI 683

RESULT 6
O90WT9 PRELIMINARY: PRT: 287 AA.
AC 090WT9:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE TNF-related apoptosis inducing ligand.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
NCBI_TaxID=9031;
RX MEDLINE-20196006; PubMed-10731132;
RC STRAIN-BERKELEY;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner K., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

```

RA Bridgman J.T., Johnson A.L.;  
 RT "TNF-related apoptosis inducing ligand (TRAIL) expression in the hen  
 RL ovary";  
 DR EMBL; AY057941; AAL23702.1;  
 DR InterPro: IPR003636; TNF-abc.  
 DR InterPro: IPR004478; TNF\_family.  
 DR Pfam: PF00229; TNF; 1.  
 DR ProDom: PD002012; TNF-abc; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 SQ SEQUENCE 287 AA; 32092 MW; DB06E1C95087B108 CRC64;

Query Match 7.7%; Score 89; DB 13; Length 287;  
 Best Local Similarity 23.4%; Pred. No. 2.3;  
 Matches 40; Conservative 24; Mismatches 65; Indels 42; Gaps 7;

OY 84 AAHVYHPR-PQDGAQAGVDG-----TVSGWEETKINS-SPLRYDROIGFTVIRA 134  
 DB 127 SAHLFRQNPADGSSRRFGLNSQSCRHATRWEDSTHSLQNTY--RDGRLRVMDA 184  
 OY 135 GLVYLCGVHFD-----EGKAVYTKLVLVNGVIALRCLEEFSAFA 175  
 DB 185 GKTYYSQITFRYSRDGAGARVSPQVQCNMKTYSQPTLLKGV-----GTK 234  
 OY 176 ASSPQPQ--LRLCQVSGLLPLRPGSSLRIRLPMHAKAPFLVYGLFQV 224  
 DB 225 CWAPAEVGHVLYQGGFLKAGDELFSVSLAIDYSDAASFTGAFRL 285

## RESULT 7

089341 PRELIMINARY; PRT; 352 AA.

AC 089341;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Matrix.  
 GN M.  
 OS Hendra virus.  
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.  
 OX NCBI\_TaxID=63330;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99058172;  
 RA Yu M., Hanson E., Langedijk J.P., Eaton B.T., Wang L.F.;  
 RT "The attachment protein of Hendra virus has high structural similarity  
 RT but limited primary sequence homology compared with viruses in the  
 RT genus Paramyxovirus";  
 RL Virology 251:227-233(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20481636; PubMed=11024125;  
 RA Wang L.F., Yu M., Hanson E., Pritchard L.I., Shiell B.,  
 RA Michalski W.P., Eaton B.T.;  
 RT "The exceptionally large genome of Hendra virus: support for creation  
 RT of a new genus within the family Paramyxoviridae";  
 RL J. Virol. 74:9972-9979(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Wang L.F., Yu M., Pritchard L.I., Hanson E., Eaton B.T.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF017149; AAC83191.1;  
 DR InterPro: IPR000982; Matrix.  
 DR Pfam: PF000661; Matrix; 1.  
 DR ProDom: PD000741; Matrix; 1.  
 SQ SEQUENCE 352 AA; 39793 MW; 79E238DE496828D5 CRC64;

Query Match 7.7%; Score 89; DB 12; Length 352;  
 Best Local Similarity 28.1%; Pred. No. 3;  
 Matches 39; Conservative 20; Mismatches 50; Indels 30; Gaps 8;

DB 197 RNNNAIAFNLVLYLKIDADLAKAGIOGSEFDK-DGTRKVAS-----PMLHGNF-VRRACKY 249  
 OY 138 YL--YCOVHFDEGKAVY-----TKLVLVNGVIALRCLEEFSAFASSPQRLQCV 188  
 DB 250 YSEYCKRKIDMKLOFISGISTGLSHIKINGVISKRLPAQM-----GLQKNLIC-- 299  
 OY 189 SGLPLRPGSSLRIRLTPM 207  
 DB 300 FSLMDINPM-----LNRLTW 314

## RESULT 8

09V762 PRELIMINARY; PRT; 224 AA.

ID 09V762;  
 AC 09V762;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE CG10205 protein.  
 GN CG10205.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY.  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers J.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Plannkoch C., Baldwin D.,  
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadelu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
 RA Jatala M., Kalush F., Kaipen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spter E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003813; AAF58200.1;  
 DR FlyBase; FBgn0033970; CG10205.  
 SQ SEQUENCE 224 AA; 25025 MW; 780B9EF7BBAF156 CRC64;

Query Match 7.6%; Score 88; DB 5; Length 224;

Best Local Similarity 22.6%; Pred. No. 2.1;  
Matches 44; Conservative 27; Mismatches 52; Indels 72; Gaps 10;

OY 10 CIGLLVVVSGSATLSAOPESOELEAED-----REPELINO--TESSQ 55  
DB 13 CIGAL-----TWASLAIEESTDIE---EDVISTANOTOKVIRVHPDLPKKTTSQ 62  
OY 56 DWPFLE-OLVPRPRS-----APGRKARPRRA-----IAHYEVHPRPGODGA 98  
DB 63 MNSPFIQVOTLRPSTAFSSQROYVDSKOMRRRPRPKLLGNNAA-----SGEDAT 115  
OY 99 QAGVDGVSGWETKINSSPLKRD-----RQIGETVIRAG----- 135  
DB 116 QPALKEELKAFPKOKLKOKOLINDARGAGELSGROYELKQVEVGQVLPVQMPGPY 175  
OY 136 -LYYLXQVHFDECK 149  
DB 176 PLYVVSXTNGREGK 190

## RESULT 9

O66760 PRELIMINARY; PRT; 353 AA.

AC Q66760; 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Matrix protein  
OS Equine morbillivirus.  
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;  
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.  
OX NCBI\_TaxID=45337;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96419932; PubMed=8822631;  
RA Gould A.R.;  
RT "Comparison of the deduced matrix and fusion protein sequences of  
RT equine morbillivirus with cognate genes of the Paramyxoviridae.";  
RL Virus Res. 43:17-31(1996).  
DR EMBL; U49404; AAB39504.1; -.  
DR InterPro; IPR000982; Matrix.  
DR Pfam; PF00661; Matrix; 1.  
DR Prodom; PD000741; Matrix; 1.  
SQ SEQUENCE 353 AA; 40142 MW; EBE67A14B0FB2AD CRC64;

Query Match 7.6%; Score 88; DB 12; Length 353;  
Best Local Similarity 28.1%; Pred. No. 3.7;  
Matches 39; Conservative 20; Mismatches 50; Indels 30; Gaps 8;

OY 78 RPRRAIAHYEVHPRPGQDGAQAGVGVSGWETKINSSPLRYDROIGETVIRAGLY 137  
DB 198 RNNNAIAINLVYLKIDDLAKAGIOGSFDR-DGTRKAS-----FMLHGNF-VRRAGRY 250  
OY 138 YL--YCVHFDEGKAVY-----LKLDLVNGVIALRCLEEFSAATSAASPPQRLCOV 188  
DB 251 YSVECKRKIRIDMKLOFSLSIGSLHIKINGVSKRLFAQM-----GFOKNLC-- 300  
OY 189 SGLPLRPGSSLRITLPM 207  
DB 301 FSLMDINPM---LNRITW 315

## RESULT 10

O16727 PRELIMINARY; PRT; 532 AA.

AC Q16727; 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Placental-like alkaline phosphatase precursor (EC 3.1.3.1).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE=COLON.  
RX MEDLINE=90124311; PubMed=2297757;  
RA Gum J.R., Hicks J.W., Sack T.L., Kim Y.S.;  
RT "Molecular Cloning of Complementary DNAs Encoding Alkaline Phosphatase  
RT in Human Colon Cancer Cells.";  
RL Cancer Res. 50:1085-1091(1990).  
DR EMBL; X53279; CAAS374.1; -.  
DR HSSP; P0634; IADC.  
DR InterPro; IPR001952; Alk\_phosphatase.  
DR Pfam; PF00245; alk\_phosphatase; 1.  
DR PRINTS; PR00113; ALKPHPTASE.  
DR Prodom; PD001868; Alk\_phosphatase; 1.  
DR SMART; SM00088; alkppc; 1.  
DR PROSITE; PS00123; ALKALINE\_PHOSPHATASE; 1.  
KW Signal.  
FT CHAIN 20 532 POTENTIAL.  
FT SIGNAL  
SQ SEQUENCE 532 AA; 57399 MW; 354B345CB19A0BC4 CRC64;

Query Match 7.6%; Score 88; DB 4; Length 532;  
Best Local Similarity 28.6%; Pred. No. 6.3;  
Matches 44; Conservative 16; Mismatches 44; Indels 50; Gaps 8;

OY 76 KAPRRRAIAHYEV---HPRG---QDGAQCYDGVSGWETKINSSPLRYDROIGEF 129  
DB 400 KARDRKA---YVLLYNGRQYVLKDGARDVTESESGSPRYROQSAVPDGETHAGED 455  
OY 130 TVIRAGLYLYLXQVHFDEGKAVYLTLDLVNGV-----LALRCLEESA----- 173  
DB 456 VAV-----FARGPQAH-----LVHGQVQGTFIAHYMAFACLEPTACDLAPS 498  
OY 174 ---TAASPSPQRLCOVSGSLPIRPGSSLRIT 204  
DB 499 AGTTDAHGPSV---VPALLPLLAGTLLILGT 528

## RESULT 11

O90Q01 PRELIMINARY; PRT; 1663 AA.

AC O90Q01; 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE KIA0929 protein (Fragment).  
GN KIA0929.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE=BRIN;  
RX MEDLINE=99246063; PubMed=10231032;  
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirosewa M.,  
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XIII.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 6:63-70(1999).  
DR EMBL; AB023146; BAA76773.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 1663 AA; 173871 MW; EC94D17086A8531E CRC64;

Query Match 7.6%; Score 88; DB 4; Length 1663;  
Best Local Similarity 23.1%; Pred. No. 27;

Matches 42; Conservative 20; Mismatches 62; Indels 58; Gaps 7;

OY 33 QBELTAEDRRREPELNPOTESQ-----DWPFLEOLVPRRSAPK----- 73  
DB 1400 QPRLPAGANRPPEPHTOVORAQAEPTSPSPVSVSMKPD-LFVSLPOTAPKQPLTV 1458

QY 74 ---GRKAPRRRAIAHVEVHPRGODGAGVDTVSGMEETKINSSPLRYDROIGFT 130  
 DB 1459 PTTSGSTPGVLPHPTFEOPAPKODS-----PHLTSORPVDWYOLKKTYP 1505  
 QY 131 VIRAGLYYL-----YCOVHEDEKAVYLLKLDLVNGVLAIRCLSEFSATAPSSPGPOLRL 185  
 DB 1506 IYVQGLLALKNDTAAVQLHFVSG-----NNVLAHRSLL-----PLSGGSPPLRI 1548  
 QY 186 CQ 187  
 DB 1549 AQ 1550

## RESULT 12

QY9Y556 PRELIMINARY: PRT: 3261 AA.  
 AC Q9Y556;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE Hypothetical 357.0 kDa protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCBL\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Rhodes S., Huckle E.;  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL096858; CAB51072.1;  
 DR InterPro; IPR000504; RNA\_rec\_mot.  
 DR Pfam; PF00076; Rfam; 2.  
 DR SMART; SM00360; RRM; 2.  
 DR PROSITE; PS50102; RRM; 2.  
 KW Hypothetical protein.  
 SQ SEQUENCE 3261 AA; 357037 MW; F8B6A645D09B66C CRC64;

Query Match 7.6%; Score 88; DB 4; Length 3261;  
 Best Local Similarity 23.1%; Pred. No. 65;  
 Matches 42; Conservative 20; Mismatches 62; Indels 58; Gaps 7;

QY 33 QEELTAEDRRPELPNPQTEESQ-----DYVPLEQLVRRRSAPK----- 73  
 DB 2998 QPRLPGAPNRPPEPHPTVOQAETGPTSPSPSVSMKPDLPVSLPTQTAPKQPLFV 3056  
 QY 74 ---GRKAPRRRAIAHVEVHPRGODGAGVDTVSGMEETKINSSPLRYDROIGFT 130  
 DB 3057 PTTSGSTPGVLPHPTFEOPAPKODS-----PHLTSORPVDWYOLKKTYP 3103  
 QY 131 VIRAGLYYL-----YCOVHEDEKAVYLLKLDLVNGVLAIRCLSEFSATAPSSPGPOLRL 185  
 DB 3104 IYVQGLLALKNDTAAVQLHFVSG-----NNVLAHRSLL-----PLSGGSPPLRI 3146  
 QY 186 CQ 187  
 DB 3147 AQ 3148

## RESULT 13

QY96T58 PRELIMINARY: PRT: 3664 AA.  
 AC Q96T58;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DE Hypothetical 21, Last annotation update)  
 DE Nuclear receptor transcription cofactor.  
 GN SHARP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCBL\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Shi Y., Evans M., Xie W., Kao H.-Y., Orendtlich P., Tsai C.-C.,  
 RA Hon M., Downs R.M.;  
 RT "SHARP, an inducible cofactor that integrates nuclear receptor  
 repression and activation."  
 RL Genes Dev. 0:0-0(2001).  
 DR EMBL; AF356524; AK52750.1;  
 DR InterPro; IPR000504; RNA\_rec\_mot.  
 DR Pfam; PF00076; Rfam; 4.  
 DR PROSITE; PS50102; RRM; 4.  
 KW Receptor.  
 SQ SEQUENCE 3664 AA; 402243 MW; 5228C585335B27B CRC64;

Query Match 7.6%; Score 88; DB 4; Length 3664;  
 Best Local Similarity 23.1%; Pred. No. 75;  
 Matches 42; Conservative 20; Mismatches 62; Indels 58; Gaps 7;

QY 33 QEELTAEDRRPELPNPQTEESQ-----DYVPLEQLVRRRSAPK----- 73  
 DB 3401 QPRLPGAPNRPPEPHPTVOQAETGPTSPSPSVSMKPDLPVSLPTQTAPKQPLFV 3459  
 QY 74 ---GRKAPRRRAIAHVEVHPRGODGAGVDTVSGMEETKINSSPLRYDROIGFT 130  
 DB 3460 PTTSGSTPGVLPHPTFEOPAPKODS-----PHLTSORPVDWYOLKKTYP 3506  
 QY 131 VIRAGLYYL-----YCOVHEDEKAVYLLKLDLVNGVLAIRCLSEFSATAPSSPGPOLRL 185  
 DB 3507 IYVQGLLALKNDTAAVQLHFVSG-----NNVLAHRSLL-----PLSGGSPPLRI 3549  
 QY 186 CQ 187  
 DB 3550 AQ 3551

## RESULT 14

QY9FTN7 PRELIMINARY: PRT: 522 AA.  
 AC Q9FTN7;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DE Putative Rer1a protein (Atreria).  
 GN P0005A05.12.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzoae; Oryza.  
 NC NCBL\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 clone: P0005A05."  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AP002863; BAB16908.1;  
 DR InterPro; IPR004932; Rer1.  
 DR Pfam; PF03248; Rer1; 1.  
 SQ SEQUENCE 522 AA; 58171 MW; AEA3EE294DFE88E7F CRC64;

Query Match 7.5%; Score 87.5; DB 10; Length 522;  
 Best Local Similarity 30.3%; Pred. No. 6.9;  
 Matches 57; Conservative 21; Mismatches 57; Indels 53; Gaps 15;

QY 70 SAPKGRKAPRRRAIAHVEVHP---RPG---ODGA-----QAGVDGYSGME-----ERK 113  
 DB 201 SSSKGRGEGFAAREYVA-----SPSGFRPGFSMDGADSGTAGAAAAAAKRTDASRAFOY 256  
 QY 114 INSSPLRYDROIGFTVIRAGLYYLVCVHDE-----GKAVYLLKLDLVNGVLAIR 166  
 DB 257 LDRSTPHATGRWIGTIAV--AAITAL--RYLVQGYIYTYTGIGITL--LNLIL--GFLSM 310  
 QY 167 CLEEFSAATASSPGPOLRLCOVSGLLPLRPGSSLR--ITLP-----WNLKLA--APPL 216

Db 311 VDEAFA-AASDGP-----ALPTGSDDEFKFFIRRLPEKFFWATKALIAFWM 360  
 QY 217 TYFGLEFQV 224  
 Db 361 TFEVFQDV 368

## RESULT 15

Q9AA15 PRELIMINARY; PRT; 670 AA.  
 AC Q9AA15;  
 DT 01-JUN-2001 (TREMBLREL. 17, Created)  
 DT 01-JUN-2001 (TREMBLREL. 17, last sequence update)  
 DT 01-MAR-2002 (TREMBLREL. 20, last annotation update)  
 DE Alpha-glucosidase, putative.  
 GN CC0796.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Usterback T., Tran K., Wolf A., Yamathavan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 RL EMBL; AE005735; AAK22781.1; -  
 DR TIGR; CC0796; -  
 KW Complete proteome.  
 SQ SEQUENCE 670 AA; 72724 MW; CBEC8FF95B9300C2 CRC64;

## Query Match

Best Local Similarity 7.5%; Score 87.5; DB 16; Length 670;

Best Local Similarity 24.0%; Pred. No. 9.5;

Matches 43; Conservative 24; Mismatches 75; Indels 37; Gaps 6;

QY 58 VFLEQLVLP-----RRSAPKGRKR-----PRRAIAHYEVHPRP----- 93  
 Db 498 LPTRLILGPMDYTPGCFRNAPKDKSQFILPTVQTRGQALAMTVYDSPLTMVSDSP 557  
 QY 94 -GODGAQAGVD---GTVSGMEETKINSSPLRYDQIGETVYI--RAGLYYLCOVHFDE 147  
 Db 558 ITYAASPAQIDFISAVPTSMDETRVLSG-----EIGQYIYARRKADMMWGAMTNEA 610  
 QY 148 GKAVYIAKLDLVNGVIALKCLEFEFSATAASSPEQRLRCOVSGLLPLRPGSSLRIRTL 206  
 Db 611 GRTVKVPLSFLNGAFSAEIRDEGAPETALKTRTQSVASKDPLTLKLAIPGGGCVIRISP 669

Search completed: May 8, 2003, 02:21:23

Job time : 27.8703 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:07:47 ; Search time 8.39882 Seconds

(without alignments)  
1111.129 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162

Sequence: 1 VLSTGLALACGLLVVSL.....PMHLKAAPFLTYFGLFOVH 225

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1162	100.0	225	1	TN12_MOUSE
2	1020	87.8	249	1	TN12_MOUSE
3	109.5	9.4	272	1	TN12_MOUSE
4	102.5	8.8	316	1	TN11_MOUSE
5	90.5	7.8	260	1	TN15_MOUSE
6	90.5	7.8	318	1	TN11_MOUSE
7	88	7.6	532	1	PPBN_HUMAN
8	87	7.5	244	1	TN12_MOUSE
9	87	7.5	261	1	TN15_MOUSE
10	87	7.5	261	1	TN15_MOUSE
11	86.5	7.4	240	1	TN14_MOUSE
12	86.5	7.4	1237	1	B3A2_MOUSE
13	86	7.4	261	1	TN15_MOUSE
14	86	7.4	261	1	TN15_MOUSE
15	85	7.3	240	1	TN15_MOUSE
16	85	7.3	1237	1	B3A2_MOUSE
17	84	7.2	278	1	TN16_MOUSE
18	83.5	7.2	241	1	TN13_MOUSE
19	83.5	7.2	250	1	TN13_MOUSE
20	82	7.1	530	1	PPB1_HUMAN
21	82	7.1	535	1	PPB1_HUMAN
22	82	7.1	535	1	PPB3_HUMAN
23	80.5	6.9	246	1	CLOC_MOUSE
24	80.5	6.9	1234	1	B3A2_MOUSE
25	79.5	6.8	197	1	TN16_MOUSE
26	79	6.8	920	1	U104_MOUSE
27	79	6.8	1584	1	U104_MOUSE
28	79	6.8	2468	1	MAPB_HUMAN
29	78.5	6.8	260	1	TN15_MOUSE
30	78.5	6.8	1164	1	PRD_MOUSE
31	78	6.7	285	1	PRD_MOUSE
32	78	6.7	763	1	APP2_HUMAN
33	77.5	6.7	201	1	TN16_MOUSE

34	77.5	6.7	261	1	TN15_MOUSE
35	77.5	6.7	787	1	NASP_HUMAN
36	77	6.6	279	1	TN16_MOUSE
37	77	6.6	817	1	NAH1_BOVIN
38	76.5	6.6	214	1	SMP_MOUSE
39	76	6.5	695	1	APP2_MOUSE
40	75.5	6.5	250	1	TN13_MOUSE
41	75.5	6.5	575	1	MIS_BOVIN
42	75.5	6.5	999	1	OXRP_MOUSE
43	75.5	6.5	1058	1	UBA1_HUMAN
44	75	6.5	1004	1	SAI2_MOUSE
45	75	6.5	3305	1	APPL_MOUSE

# ALIGNMENTS

RESULT ID	TN12_MOUSE	STANDARD	PRT	225 AA
AC	054907	09CTP2		
DT	15-JUN-2002	(Rel. 41, Created)		
DT	15-JUN-2002	(Rel. 41, Last sequence update)		
DE	Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (Fragment).			
GN	TNFSF12			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Peritoneal macrophage;			
RX	MEDLINE=98070415; PubMed=9405449;			
RA	Chicheportiche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H., Hession C., Garcia I., Browning J.L.;			
RT	"TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis."			
RL	J. Biol. Chem. 272:32401-32410(1997).			
RN	(2)			
RP	SEQUENCE OF 83-225 FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Retina;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y., Aizawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S., Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R., Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T., Fleischmann W., Gaasterland T., Gissi C., King B., Koehli H., Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J., Schriml L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T., Sakai K., Okido T., Furuno M., Aono H., Baladrelli R., Barsh G., Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F., Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M., Gustigich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H., Lyons P., Marchionni L., Mashima J., Mazzaletti J., Mombaerts P., Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N., Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F., Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Williams L., Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S., Hayashizaki Y.;			
RT	"Functional annotation of a full-length mouse cDNA collection."			
RL	Nature 409:685-690(2001).			
CC	-1- FUNCTION: Binds to FN14 and possibly also to TNFSF12/APP3. Weak inducer of apoptosis in some cell types. Promotes angiogenesis and the proliferation of endothelial cells. Mediates NF-kappaB activation (by similarity).			
CC	-1- SUBUNIT: Homotrimer (Potential).			
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By similarity).			
CC	-1- TISSUE SPECIFICITY: WIDELY EXPRESSED.			
CC	-1- PTM: The soluble form is produced from the membrane form by			

Proteolytic processing (By similarity).  
 -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: AF030100; AAC53517.1; -  
 CC EMBL: AK020909; BAB32249.1; -  
 CC MGD: MGI:1196259; Tnfstf12.  
 CC InterPro: IPR000478; TNF\_family.  
 CC Pfam: PF00229; TNF; 1.  
 CC SMART: SM00207; TNF; 1.  
 CC PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
 CC PROSITE: PS50049; TNF\_2; 1.  
 CC Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.  
 CC NON\_TER 1  
 CC CHAIN <1 225  
 CC FT CHAIN 70 225  
 CC FT CHAIN 70 225  
 CC FT TRANSMEM <1 21  
 CC FT DOMAIN 22 225  
 CC FT SITE 69 70  
 CC FT DISULFID 167 186  
 CC FT CARBOHYD 115 115  
 CC SEQUENCE 225 AA; 24781 MW; 90C412CC0480659B CRC64;  
 CC  
 CC Query Match 100.0%; Score 1162; DB 1; Length 225;  
 CC Best Local Similarity 100.0%; Pred. No. 4.3e-97;  
 CC Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC 1 VLSIGLALACGLLLVYVSLGSMATLSAEPSEELTAEDRREPPELNQTEESDYVPE 60  
 CC DB 1 VLSIGLALACGLLLVYVSLGSMATLSAEPSEELTAEDRREPPELNQTEESDYVPE 60  
 CC QY 61 LEQVLRPRRSAPKGRKAPRAIAAHYEHVPRPGDGAQAGVDGTVSGWEETKINS 120  
 CC DB 61 LEQVLRPRRSAPKGRKAPRAIAAHYEHVPRPGDGAQAGVDGTVSGWEETKINS 120  
 CC QY 121 RYDRQIGFETVIRAGLYLYCOVHFDEGAVYIKLIDLVNGVLAARCLSEPSATAAS 180  
 CC DB 121 RYDRQIGFETVIRAGLYLYCOVHFDEGAVYIKLIDLVNGVLAARCLSEPSATAAS 180  
 CC QY 181 POLRLQVSGLLPLRPGSSLRIRLPLMAHLKAAPLTFYFGLFQVH 225  
 CC DB 181 POLRLQVSGLLPLRPGSSLRIRLPLMAHLKAAPLTFYFGLFQVH 225

RESULT 2  
 TN12\_HUMAN STANDARD; PRT; 249 AA.  
 AC 043508; Q8WU27;  
 DT 15-JUN-2002 (Ref. 41, Created)  
 DT 15-JUN-2002 (Ref. 41, Last sequence update)  
 DE 15-JUN-2002 (Ref. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (APO3 ligand).  
 GN TNFSF12 OR APO3L OR DR3LG.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND N-TERMINUS OF SOLUBLE FORM.  
 RC TISSUE-Tonsil, and Fetal liver;  
 RX MEDLINE=98070415; PubMed=9405449;  
 RA Chicheportliche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,

RA Hession C., Garcia I., Browning J.L.;  
 RT "TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis."  
 RL J. Biol. Chem. 272:32401-32410(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal Kidney;  
 RX MEDLINE=98228355; PubMed=9560343;  
 RA Masters S.A., Sheridan J.P., Pitti R.M., Brush J., Goddard A., Ashkenazi A.;  
 RT "Identification of a ligand for the death-domain-containing receptor APO3."  
 RL Curr. Biol. 8:525-528(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Tonsil;  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP FUNCTION.  
 RA PubMed=10085077;  
 RX Lynch C.N., Wang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S.R.;  
 RT "TWEAK induces angiogenesis and proliferation of endothelial cells."  
 RL J. Biol. Chem. 274:8455-8459(1999).  
 CC -1- FUNCTION: Binds to TNFSF12/APO3. Weak inducer of apoptosis in some cell types. Mediates NF-kappaB activation. May promote angiogenesis and the proliferation of endothelial cells.  
 CC -1- SUBUNIT: HOMOTRIMER (POTENTIAL).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted. skeletal muscle, brain, colon, small intestine, lung, ovary, prostate, spleen, lymph node, appendix and peripheral blood lymphocytes. Low expression in kidney, testis, liver, placenta, thymus and bone marrow. Also detected in fetal kidney, liver, lung and brain.  
 CC -1- PM: The soluble form derives from the membrane form by proteolytic processing.  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to a frameshift in position 125.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: AF030099; AAC51923.1; -  
 CC EMBL: AF035872; AAC39724.1; -  
 CC EMBL: BC019047; AAH19047.1; ALT\_FRAME.  
 CC Genew: HGNC:11927; TNFSF12.  
 CC MIM: 602695; -  
 CC InterPro: IPR000478; TNF\_family.  
 CC Pfam: PF00229; TNF; 1.  
 CC SMART: SM00207; TNF; 1.  
 CC PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
 CC PROSITE: PS50049; TNF\_2; 1.  
 CC Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.  
 CC CHAIN 1 249  
 CC FT CHAIN 94 249  
 CC FT CHAIN 94 249  
 CC FT DOMAIN 1 21  
 CC FT TRANSMEM 22 42  
 CC FT TRANSMEM 22 42  
 CC FT DOMAIN 43 249  
 CC FT SITE 93 94  
 CC FT CARBOHYD 139 139  
 CC SEQUENCE 249 AA; 27216 MW; E660843361C28EBA CRC64;

Query Match 87.8%; Score 1020; DB 1; Length 249;  
 Best Local Similarity 88.8%; Pred. No. 2,6e-84;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY 2 LSLGLALACGLLVVSGSWATLSAQEPSELTAEEDRREPPELNPOTEEQDVVPL 61  
 DB 26 LGLGLALACGLLVVSGSWATLSAQEPSELTAEEDRREPPELNPOTEEQDVVPL 85  
 QY 62 EOLVPRRSPKGRKAPRAIAHYEVHPRGDGAQAGVDTGSGWEETKINSSPLR 121  
 DB 86 NRIYVRRRSPKGRKTRARAIAMHVEHPRGQDGAQAGVDTGSGWEETKINSSPLR 145  
 QY 122 YDRQIGFTYIRAGLYLYCQVHFDEGKAVYIKLDLVNGVIALRCLEEFSSATPAASP 181  
 DB 146 YNRQIGFTYIRAGLYLYCQVHFDEGKAVYIKLDLVNGVIALRCLEEFSSATPAASP 205  
 QY 182 QLRICQVSGILPLRPSSSLRIRLTPMAHLKAPFLTYFGFLFOVH 225  
 DB 206 QLRICQVSGILPLRPSSSLRIRLTPMAHLKAPFLTYFGFLFOVH 249

RESULT 3  
 TNF5\_CHICK STANDARD; PRT; 272 AA.  
 ID TNF5\_CHICK Q918D8;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-l) (CD154 protein).  
 GN TNF5F5 OR CD40LG OR CD40L.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=White Leghorn; TISSUE=Spleen;  
 RA Tregaskes C.A., Young J.R., Burnside J.;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBD databases.  
 CC -1- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell proliferation in the presence of IL-4. Involved in immunoglobulin class switching (By similarity).  
 CC -1- SUBUNIT: Homotrimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an extracellular soluble form (By similarity).  
 CC -1- PTM: The soluble form derives from the membrane form by proteolytic processing (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AJ243435; CAB95748.1; -  
 DR HSSP: P29965; IALY.  
 DR InterPro: IPR003263; TNF\_5.  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR00478; TNF\_family.  
 DR Pfam: PF00229; TNF; 1.  
 DR ProDom: PDD02012; TNF\_abc; 1.  
 DR ProDom: PDD08600; TNF\_5; 1.  
 DR SMART: SMO0207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 KW Cytokine; Transmembrane; Glycoprotein; Signal-anchor.  
 FT CHAIN 1 272 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY

FT CHAIN 111 272  
 FT DOMAIN 1 23  
 FT TRANSMEM 24 44  
 FT DOMAIN 45 272  
 FT SITE 110 111  
 FT DISULFID 190 229  
 FT CARBOHYD 124 124  
 FT CARBOHYD 146 146  
 FT CARBOHYD 251 251  
 SQ SEQUENCE 272 AA; 30862 MW; 5409F24A8E53CCD7 CRC64;

Query Match 9.4%; Score 109.5; DB 1; Length 272;  
 Best Local Similarity 22.5%; Pred. No. 0.012; Indels 53; Gaps 10;  
 Matches 58; Conservative 40; Mismatches 107;

QY 1 VLSGLALACGLLVVSGSWATLSAQEPSELTAEEDRREPPELNPOTEEQDVVPL 45  
 DB 34 VQIGIVLFCVLYHMKMDKMEVLSLNDYIFLRKQKQKQSGEDKSTLLDCEKVLKGRQ 93  
 QY 46 ELNPOTEEQDVVPLLEQLVPRRSPKGRK-----APRRATAIHYEVHPRGDGAQA 100  
 DB 94 DLCKDKRTASEELPKFEMHRSHEHPILKSNETSVAEERKOPATHTLA-----GVKS 145  
 QY 101 GVDGTVSGWEETK-INSSPLRYDRDIGFTYIRAGLYLYCQVHFDEGKA-----V 151  
 DB 146 NTVVRVLRKMTTSYAPDTSLISYH--GRIKYEKALYYIYVSVCYTAASAPPTLYI 203  
 QY 152 YIKLDLVNGVIALRCLEEFSSATPAASPQRLICQV-----SGLPLRPGSSLRIRLTP 206  
 DB 204 YLYLPMEDERL-MKLDLDHSTSTA-----LCLEQSTREGVEELKQGMVFNVTVD 254  
 QY 207 WAHLKAPFLTYFGFLFOV 224  
 DB 255 STAVNVNPGNTYFGMKL 272

RESULT 4  
 TN11\_MOUSE STANDARD; PRT; 316 AA.  
 ID TN11\_MOUSE Q35235;  
 AC Q35235; Q35306; Q9J1Y0; Q9J1K9;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 11 (Receptor activator of nuclear factor kappa B ligand) (RANKL) (TNF-related activation induced cytokine) (TRANCE) (Osteoprotegerin ligand) (OPGL) (Osteoclast differentiation factor) (OPF) (Osteoclastogenesis-inhibitory factor) (OCIF).  
 GN TNF5F11 OR RANKL OR TRANCE OR OPGL.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Hydrinoma;  
 RA MEDLINE-97460112; PubMed-9312132;  
 RA Wong B.R., Rho J., Arron J., Robinson E., Orlinick J., Chao M., Kalachkov S., Cayant E., Bartlett F.S. III, Frankel W.N., Lee S.Y., Choi Y.;  
 RA "TRANCE is a novel ligand of the tumor necrosis factor receptor family that activates c-Jun N-terminal kinase in T cells";  
 RT J. Biol. Chem. 272:25190-25194(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Thymic lymphoma;  
 RA MEDLINE-98032977; PubMed-9367155;  
 RA Anderson D.M., Maraskovsky E., Billingsley W.L., Dougan W.C., Tometsko M.E., Roux E.R., Teepe M.C., Dubose R.F., Cosman D., Galibert L.;

RT "A homologue of the TNF receptor and its ligand enhance T-cell growth  
and dendritic-cell function." ;  
RL Nature 390:175-179(1997).  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE-Bone marrow;  
RX MEDLINE-98227661; Pubmed-9568710;  
RA Lacey D.L., Timsa E., Tan H.-L., Kelley M.J., Dunstan C.R.,  
BA Burgess T., Elliott R., Colombero A., Elliott G., Scully S., Hu H.,  
RT Sullivan V., Hawkins N., Davy E., Capprelli C., Eli A., Qian Y.-X.,  
RA Boyle W.J., Sarosi I., Shalhoub V., Senaldi G., Guo Y., Delaney J.,  
RT "Osteoprotegerin ligand is a cytokine that regulates osteoclast  
differentiation and activation." ;  
RL Cell 93:165-176(1998).  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE-Bone marrow stroma;  
RX MEDLINE-98188248; Pubmed-9520411;  
RA Yasuda H., Shima N., Nakagawa N., Yamaguchi K., Kinoshita M.,  
BA Mohtazaki S.-I., Tomoyasu A., Iano K., Goto M., Murakami A.,  
RT Morinaga T., Higashio K., Udagawa N., Takahashi N., Suda T.;  
RT "Osteoclast differentiation factor is a ligand for  
osteoprotegerin/osteoclastogenesis-inhibitory factor and is identical  
to TRANCE/RANKL." ;  
RL Proc. Natl. Acad. Sci. U.S.A. 95:3597-3602(1998).  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC STRAIN-129;  
RX MEDLINE-99214075; Pubmed-10196481;  
RA Kodaira K., Kodaira K., Mizuno A., Yasuda H., Shima N., Murakami A.,  
RT Ueda M., Higashio K.;  
RT "Cloning and characterization of the gene encoding mouse osteoclast  
differentiation factor." ;  
RL Gene 230:121-127(1999).  
RN [6]  
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).  
RX MEDLINE-21150053; Pubmed-11250921;  
RA Ikeda T., Kasai M., Utsuyama M., Hirokawa K.;  
RT "Determination of three isoforms of the receptor activator of nuclear  
factor-kappaB ligand and their differential expression in bone and  
thymus." ;  
RL Endocrinology 142:1419-1426(2001).  
RN [7]  
RP SEQUENCE OF 139-147, PROCESSING, AND N-GLYCOSYLATION.  
RX Pubmed-10224132;  
RA Lum L., Wong B.R., Josien R., Becherer J.D., Erdjument-Bromage H.,  
RT Schlegendorff J., Tempst P., Choi Y., Blobel C.P.;  
RT "Evidence for a role of a tumor necrosis factor-alpha  
(TNF-alpha)-converting enzyme-like protease in shedding of TRANCE, a  
TNF family member involved in osteoclastogenesis and dendritic cell  
survival." ;  
RL J. Biol. Chem. 274:13613-13618(1999).  
RN [8]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 158-316.  
RX Pubmed-11581298;  
RA Lam J., Nelson C.A., Rose F.P., Tietelbaum S.L., Fremont D.H.;  
RT "Crystal structure of the TRANCE/RANKL cytokine reveals determinants  
of receptor-ligand specificity." ;  
RL J. Clin. Invest. 108:971-979(2001).  
RN [9]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 137-316.  
RX Pubmed-11733492;  
RA Ito S., Wakabayashi K., Ubukata O., Hayashi S., Okada F., Hata T.;  
RT "Crystal structure of the extracellular domain of mouse RANK ligand at  
2.2-A resolution." ;  
RL J. Biol. Chem. 277:6631-6636(2002).  
RN -1-  
RP FUNCTION: Cytokine that binds to TNFRSF11B/OPG and to  
TNFRSF11A/RANK. Osteoclast differentiation and activation factor.  
CC Augments the ability of dendritic cells to stimulate naive T-cell  
proliferation. May be an important regulator of interactions  
between T cells and dendritic cells and may play a role in the  
regulation of the T cell-dependent immune response. May also play



CC -1- SUBUNIT: Homotrimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By similarity).  
 CC -1- TISSUE SPECIFICITY: Highly expressed in thymus and bone tissues.  
 CC -1- PTM: The soluble form derives from the membrane form by proteolytic processing (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AF187319; AAG17031.1; -  
 CC EMBL: AF425669; AAL23963.1; -  
 CC HSP: P50591; 1D0G  
 CC InterPro: IPR003636; TNF\_abC.  
 CC InterPro: IPR000478; TNF\_family.  
 CC Pfam: PF00229; TNF; 1.  
 CC ProDom: PD002012; TNF\_abC; 1.  
 CC SMART: SM00207; TNF; 1.  
 CC PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
 CC PROSITE: PS50049; TNF\_2; 1.  
 CC CytoKine: Differentiation; Receptor; glycoprotein; Transmembrane; Signal-anchor.  
 CC CHAIN 1 318 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 CC CHAIN 141 318 MEMBER 11. MEMBRANE FORM.  
 CC TRANSMEM 48 68 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 CC TRANSMEM 48 68 MEMBER 11. SOLUBLE FORM.  
 CC DOMAIN 69 318 CYTOPLASMIC (POTENTIAL).  
 CC SITE 140 141 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN) (POTENTIAL).  
 CC CARBOHYD 199 199 EXTRACELLULAR (POTENTIAL).  
 CC CARBOHYD 264 264 CLEAVAGE (BY SIMILARITY).  
 CC CONFLICT 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC CONFLICT 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CC SEQUENCE 318 AA; 35370 MW; 4B87A4D706AD098F CRC64;  
 SO  
 Query Match 7.88; Score 90.5; DB 1; Length 318;  
 Best Local Similarity 23.38; Pred. No. 0.72;  
 Matches 64; Conservative 36; Mismatches 108; Indels 67; Gaps 12;

DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Alkaline phosphatase, placental-like precursor (PC 3.1.3.1) (Nagao isozyme) (Germ-cell alkaline phosphatase) (PLAP-like) (ALP-1).  
 GN ALPPL2 OR ALPPL.  
 OS Homo sapiens (human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90283879; PubMed=2162249;  
 RA Lowe M.E., Strauss A.W.;  
 RT "Expression of a Nagao-type, phosphatidylinositol-glycan anchored RT alkaline phosphatase in human choriocarcinomas";  
 RL Cancer Res. 50:3956-3962(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Choriocarcinoma;  
 RX MEDLINE=89308696; PubMed=2745460;  
 RA Matanabe S., Matanabe T., Li W.L., Soong B.-W., Chou J.Y.;  
 RT "Expression of the germ cell alkaline phosphatase gene in human RT choriocarcinoma cells";  
 RL J. Biol. Chem. 264:12611-12619(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88203632; PubMed=2834730;  
 RA Millan J.L., Manes T.;  
 RT "Seminoma-derived Nagao isozyme is encoded by a germ-cell alkaline phosphatase gene";  
 RL Proc. Natl. Acad. Sci. U.S.A. 85:3024-3028(1988).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Placenta;  
 RA Strausberg R.;  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE OF 1-157 FROM N.A.  
 RX MEDLINE=88262578; PubMed=387245;  
 RA Shen L.P., Liu H., Kan Y.W., Kam W.;  
 RT "5' nucleotide sequence of a putative human placental alkaline phosphatase-like gene";  
 RL Nucleic Acids Res. 16:5694-5694(1988).  
 CC -1- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an alcohol + phosphate.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
 CC -1- TISSUE SPECIFICITY: TRACE AMOUNTS IN THE TESTIS AND THYMUS, AND IN ELEVATED AMOUNTS IN GERM CELL TUMORS.  
 CC -1- MISCELLANEOUS: IN MOST MAMMALS THERE ARE FOUR DIFFERENT ISOZYMES: PLACENTAL, PLACENTAL-LIKE, INTESTINAL AND TISSUE NON-SPECIFIC (LIVER/BONE/ KIDNEY).  
 CC -1- SIMILARITY: BELONGS TO THE ALKALINE PHOSPHATASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X55958; CA39425.1; -  
 CC EMBL: J04948; AAA51700.1; -  
 CC EMBL: J03252; AAA98616.1; -  
 CC EMBL: BC014139; AAH14139.1; -  
 CC EMBL: X07247; CA30232.1; ALU\_SEQ.  
 CC PIR: A31336; A31336.  
 CC PIR: A34333; A34333.  
 CC PIR: S00974; S00974.  
 CC HSP: P00634; 1AJC.  
 CC SlenA-2DPAGE: P10696; -  
 CC Genew: HGNC:441; ALPPL2.  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)

DR InterPro: IPR001952; Alk-phosphatase.  
 DR Pfam: PF00245; alk-phosphatase; 1.  
 DR PRINTS: PR00113; ALKPHPTASE.  
 DR PRODOM: PD001868; Alk-phosphatase; 1.  
 DR SMART: SM00098; alkppc; 1.  
 DR PROSITE: PS00123; ALKALINE\_PHOSPHATASE; 1.  
 DR HydroLase; Zinc; Magnesium; Phosphorylation; Transmembrane;  
 KW Multigene family; Glycoprotein; GPI-anchor; Signal.  
 FT SIGNAL 1  
 FT CHAIN 19 POTENTIAL  
 FT PROPEP 20 503 ALKALINE PHOSPHATASE, PLACENTAL-LIKE.  
 FT ACCT\_SITE 111 111 REMOVED IN MATURE FORM.  
 FT MOI\_SITE 111 111 PHOSPHORYLATION.  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT CONFLICT 268 268 I -> M (IN REF. 2, 4 AND 5).  
 FT CONFLICT 57 57 I -> M (IN REF. 3).  
 FT CONFLICT 152 152 V -> M (IN REF. 2, 4 AND 5).  
 FT CONFLICT 178 178 T -> A (IN REF. 4).  
 FT CONFLICT 260 260 H -> R (IN REF. 4).  
 FT CONFLICT 273 273 L -> M (IN REF. 4).  
 FT CONFLICT 316 316 L -> R (IN REF. 1 AND 4).  
 FT CONFLICT 380 380 V -> L (IN REF. 2).  
 FT CONFLICT 498 498 P -> R (IN REF. 2 AND 4).  
 SQ SEQUENCE 532 AA; 57315 MM; 84AB5B28F13D6D82 CRC64;  
 Query Match 7.6%; Score 88; DB 1; Length 532;  
 Best Local Similarity 28.6%; Pred. No. 2.2;  
 Matches 44; Conservative 16; Mismatches 44; Indels 50; Gaps 8;  
 QY 76 KARPRAIAHYEV---HPRPG---QDGAQAGVDGTSGWETKINSSPLRYDQIGEF 129  
 DB 400 KARDKRA---YVLLYNGPGYVLKDGARPDVTESSGSPRYQSNVPLDGERHAGD 455  
 QY 130 TVIRAGLYYLICQVHFEDGKAVYLYKLDLYNGV-----LAKRLEESA----- 173  
 DB 456 VAV-----FANGPOAH-----LVHGVQRFITAHYMAFACLEPTACDLAPP 498  
 QY 174 ---TAASSPGQRLRCQVSGILPLRPGSSLRIT 204  
 DB 499 AGTDAHPGPSV-----VPALLPLLAGTLLLLGT 528  
 RESULT 8  
 TNFC\_HUMAN STANDARD; PRT; 244 AA.  
 ID 006643; P78370; Q99761;  
 AC 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lymphotoxin-beta (LFB-beta) (Tumor necrosis factor C) (TNF-C) (Tumor  
 DE necrosis factor ligand superfamily member 3).  
 GN LFB OR TNFSF3 OR TNFC.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OC NCBI\_TaxID:9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.  
 RC TISSUE=T-cell;  
 RX MEDLINE=93208881; PubMed=7916655;  
 RA Browning J.L., Ngam-Ek A., Lawton P., Demarinis J., Tizard R.,  
 RA Chow E.P., Hessson C., O'Brine-Greco B., Foley S.F., Ware C.F.;  
 RT "Lymphotoxin beta, a novel member of the TNF family that forms a  
 RT heteromeric complex with lymphotoxin on the cell surface.";  
 RL Cell 72:847-856(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RX MEDLINE=97445965; PubMed=9299492;  
 RA Marzocha K., Renard N., Charlot C., Bienvenu J., Colflier B.,  
 RA Salles G.;  
 RT "Identification of two lymphotoxin beta isoforms expressed in human  
 RT lymphoid cell lines and non-Hodgkin's lymphomas.";  
 RL Biochem. Biophys. Res. Commun. 238:273-276(1997).

[3]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Neville M.J., Milner C.M., Campbell R.D.;  
 RT "A new member of the immunoglobulin superfamily and a V-ATPase G  
 RT subunit are amongst the predicted products of novel genes close to the  
 RT TNF locus in the human MHC.";  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,  
 RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,  
 RA Lasky S., Hood L.;  
 RT "Sequence of the human major histocompatibility complex class III  
 RT region.";  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Shina S., Tamiya G., Oka A., Inoko H.;  
 RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS GLU-70 AND  
 RP PRO-111.  
 RA Rieder M.J., Arnel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,  
 RA Poel C.L., Roth E.J., Yi O., Nickerson D.A.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 CC - FUNCTION: Cytokine that binds to LTRR/TNFRSF3. May play a specific  
 CC role in immune response regulation. Provides the membrane anchor  
 CC for the attachment of the heterotrimeric complex to the cell  
 CC surface. Isoform 2 is probably non-functional.  
 CC - SUBUNIT: Heterotrimer of either two LTB and one LTA subunits or  
 CC (less prevalent) one LTB and two LTA subunits.  
 CC - SUBCELLULAR LOCATION: Type II membrane protein (potential).  
 CC - ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are  
 CC produced by alternative splicing.  
 CC - TISSUE SPECIFICITY: SPLEEN AND THYMUS.  
 CC - SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL: L11016; AAA99888.1; -  
 DR EMBL: U89922; AAC51769.1; -  
 DR EMBL: U79029; AAB37342.1; -  
 DR EMBL: L11015; AAA36191.1; -  
 DR EMBL: Y14768; CAA75069.1; -  
 DR EMBL: AF129756; AAD18089.1; -  
 DR EMBL: AP000505; BAB63395.1; -  
 DR EMBL: AY070219; AAL49954.1; -  
 DR EMBL: AY070219; AAL49955.1; -  
 DR PIR: A46066; A46066.  
 DR HSSP: P01374; 1TNR.  
 DR GeneW: HGNC:6711; LTB.  
 DR MIM: 600978; -  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR000478; TNF\_family.  
 DR Pfam: PF00229; TNF; 1.  
 DR PRINTS: PR01234; TNCRSISRCT.  
 DR PRODOM: PD002012; TNF\_abc; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS00049; TNF\_2; 1.  
 KW Cytokine; Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor;  
 FT Alternative splicing; Polymorphism.  
 FT DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 19 48 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT TRANSMEM 49 222 EXTRACELLULAR (POTENTIAL).

FT	CAROHYD	222	222	N-LINKED (GLCNAG.) (POTENTIAL)
FT	VARSPLIC	53	77	GLNETHAPRQADNOGEGFOLPEE -> GLGFRCQRSSQ
FT	VARSPLIC	78	244	KOISAPGSOLFPS (IN ISOFORM 2).
FT	VARIANT	70	70	MISSING (IN ISOFORM 2).
FT	VARIANT	111	111	/FTID-VAR_013025.
FT	CONFLICT	60	69	A -> P. /FTID-VAR_013026.
FT	SEQUENCE	244 AA:	25390 MM;	DPGAAOQGL -> GLSAPSGGRT (IN REF. 2; AAB37342).
SQ	Query Match	7.5%;	Score 87;	DB 1;
	Best Local Similarity	23.0%;	Pred. NO. 1.1;	Length 244;
Matches	64;	Conservative	26;	Mismatches 80; Indels 108; Gaps 13.
Oy	3 SIGLALA---	CIGLLLVVSLGSMATLSAQESQEELTAEDR-----	-REPPEL 47	
Db	18 SLLAVLAGTSLVTILLNAVPIVLVALVYPDGGIVLETETDRPDQAQGLGFOKLPKE	77		
Oy	48 NPQTESQDVVPFLQLVPRRSAPRKRAFRRAIAAHYEYHPRPGDGAQGVDTVS	107		
Db	78 ERET---DLSPGLP-	102		
Oy	108 GREETKINS--SSPLRYDRQIGETVIYRAGLYLYLCQVHF-----	-DEGRAVYLKL 155		
Db	103 GHETTKEQAFLETSGFQSDAES-LALPDQGLTYLYLVGRGRAPRGGDGPGRSVTLKS	161		
Oy	156 DLVNGVIALRCLEBSATAASSPG-PQLRL-----	-CQVS 189		
Db	162 SL-----YRAGAYGCPTEPELLLEGAFETVYPDIPAROGVGPMYSVGRC	208		
Oy	190 GILPAPRSSLRIRTPNAHLKAAPFL--TYFGLFQV	224		
Db	209 GLVOLRRGERVVNI--SHPDWDFARKTKFFGAVMV	243		
RESULT 9				
TNF5_AOTTR	TNF5_AOTTR	STANDARD;	PRT;	261 AA.
AC	Q9BDM3;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
De	Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-L)			
Gn	(CD134 protein).			
OS	TNFSF5 OR CD40LG OR CD40L.			
OC	Acinus trivirgatus (Night monkey) (Douroucoulil).			
CC	Eumariota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Acutus.			
OX	NCBI_TaxID=9505;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Lymphocytes;			
RX	MEDLINE=21383618; PubMed=11491535;			
RA	Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P., Weiss M.R., Ansatl A.A.;			
RT	"Cloning, sequencing, and homology analysis of nonhuman primate Fas/Fas-ligand and co-stimulatory molecules.";			
RT	Immunogenetics 53:315-328(2001).			
RL	-1- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell proliferation in the absence of co-stimulus as well as IgE production in the presence of IL-4. Involved in immunoglobulin class switching (By similarity).			
CC	-1- SUBUNIT: HOMODIMER (By similarity).			
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an extracellular soluble form (By similarity).			
CC	-1- PTM: The soluble form derives from the membrane form by proteolytic processing (By similarity).			
CC	-1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation -			

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF344660; AKR37542.1; -
DR HSSP; P29965; 1ALY.
DR InterPro; IPR003263; TNF_5.
DR InterPro; IPR003636; TNF_abcd.
DR InterPro; IPR000478; TNF_family.
DR Pfam; PF00229; TNF_1.
DR ProDom; PD002012; TNF_abc; 1.
DR ProDom; PD008600; TNF_5; 1.
DR SMART; SM00207; TNF_1.
DR PROSITE; PS00251; TNF_1; 1.
DR PROSITE; PS50049; TNF_2; 1.
KW Cytochrome; Transmembrane; Glycoprotein; Signal-anchor.
FT FT 1 261
FT CHAIN
FT 113 261
FT DOMAIN 1 22
FT TRANSSEM 23 43
FT FT
FT DOMAIN 44 261
FT SITE 112 113
FT DISULFID 178 218
FT CARBOHYD 240 240
SQ SEQUENCE 261 AA; 29357 MW; 85E1588B507901B5 CRC64;

Query Match 7.58; Score 87; DB 1; Length 261;
Best Local Similarity 25.98; Pred. No. 1.2;
Matches 35; Conservative 21; Mismatches 55; Indels 24; Gaps 6;

QY 105 TVSGME-----TKNSSPLRYDROIGFTYRAGLYLYCQVHFDESKAYTLKLDLYN 160
Db 136 SVLQMAEKGYYTMSNNLVLTENGKQL--TVKRGGLYIYAQVTFCSNREASQAPLAS 192
QY 161 GVALRLCLEEF-----SATPASSPGPOLRLC-----QVSGILPLRPSSIRITLPMW 209
Db 193 --LCLKPNNRERILLRANRAHSSAKP-----CGQGSIHGSIPELQPGASVFWVTDSQ 246
QY 210 LKAAPFTYFGGLFOV 224
Db 247 VSHGTGFTSGFLTKL 261

RESULT 10
TNF5_CALJA STANDARD; PRT; 261 AA.
AC OSBDN3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-
DE l) (CD154 protein).
GN TNFSF5 OR CD40LG OR CD40L.
OS Callithrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callithrix.
OC NCBI_TaxID=9483;
OX 11
RN RP SEQUENCE FROM N.A.
RC TISSUE=Lymphocytes;
RX MEDLINE=21363618; PubMed=11491535;
RA Villinger F., Bostlik P., Mayne A.E., King C.T., Genain C.P.,
RA Weiss W.R., Ansari A.A.;
RT "Cloning, sequencing, and homology analysis of nonhuman primate
RT Fas/FasL ligand and co-stimulatory molecules."
RT Immunogenetics 53:315-328(2001).
RL -!- FUNCTION: Cytochrome that binds to TNFRSF5. Mediates B-cell

```



proliferation in the absence of co-stimulus as well as IGE production in the presence of IL-4. Involved in immunoglobulin class switching (By similarity).

-1- SUBUNIT: HOMOTRIMER (By similarity).

-1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an extracellular soluble form (By similarity).

-1- PFM: The soluble form derives from the membrane form by proteolytic processing (By similarity).

-1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL: AF344844; AAK37603.1; -

DR HSSP: P29965; 1ALY.

DR InterPro: IPR003263; TNF\_5.

DR InterPro: IPR003636; TNF\_abc.

DR InterPro: IPR00478; TNF\_family.

DR Pfam: PF00229; TNF\_1.

DR ProDom: PD002012; TNF\_abc; 1.

DR ProDom: PD008600; TNF\_5; 1.

DR SMART: SM00251; TNF\_1; 1.

DR PROSITE: PS00493; TNF\_2; 1.

DR PROSITE: PS50049; TNF\_2; 1.

RW Cytokine; Transmembrane; Glycoprotein; Signal-anchor.

FT CHAIN 1 261

FT TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY

FT MEMBER 5, MEMBRANE FORM.

FT TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY

FT MEMBER 5, SOLUBLE FORM (BY SIMILARITY).

FT CYTOPLASMIC (POTENTIAL).

FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)

FT (POTENTIAL).

FT EXTRACELLULAR (POTENTIAL).

FT SITE 112 113

FT CLEAVAGE (BY SIMILARITY).

FT DISULFID 178 218

FT CARBOHD 240 240

FT N-LINKED (GLCNAC. . .) (POTENTIAL).

SO SEQUENCE 261 AA; 29360 MW; 10CA58D923754EB CRC64;

Query Match 7.5%; Score 87; DB 1; Length 261;

Best Local Similarity 25.9%; Pred. No. 1.2;

Matches 35; Conservative 21; Mismatches 55; Indels 24; Gaps 6;

OY 105 TVSGME---TKINSSPLRYDRIQGEFTIRAGLYLYCOVHFDEGRAVYLKLDLVN 160

DB 136 SVIOWAEKGYTSMNNVLENGKQL--TVKRGGLYTYAQTFCNSNEASSQAPFTAS 192

OY 161 GVLAIRCEEF-----SATASSPGPOLRIC-----QVSGLLPLRPSSGLRIRLPMAN 209

DB 193 -LICKPNNRERILLRANHTSSAKP---CGQOSIHILGFIELDPGASVFNVTDSQ 246

OY 210 LKAAPFLVYFGFVOV 224

DB 247 VSHGTGTFSGFLKL 261

RESULT 11

TN14\_HUMAN STANDARD: PRT: 240 AA.

AC Q43557; Q75476; Q96UD2; Q8WV98;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Tumor necrosis factor ligand superfamily member 14 (Herpesvirus entry mediator-ligand) (HYEM-L).

DE TNFSF14 OR LIGHT OR HYEM-L.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI\_TaxID=9606;

[1]

SEQUENCE FROM N.A. (ISOFORM 1).

MEDLINE=96122340; PubMed=9462508;

MAUI D.N., Eder R., Montgomery R.I., Kochei K.D., Cheung T.C., Yu G.-L., Ruben S., Murphy M., Eisenberg R.J., Cohen G.H., Spear P.G., Ware C.F.;

"LIGHT", a new member of the TNF superfamily, and lymphotoxin alpha are ligands for herpesvirus entry mediator.";

Immunity 8:21-30(1998).

[2]

SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.

TISSUE=Liver:

MEDLINE=96438532; PubMed=9765287;

Hartop J.A., McDonnell P.C., Brigham-Burke M., Lyn S.D., Minton J., Tan K.B., Dede K., Spananato J., Silverman C., Hensley P., DiPietro R., Emery J.G., Deen K., Eichman C., Chabot-Fletcher M., Truneh A., Young P.R.;

"Herpesvirus entry mediator ligand (HYEM-L): a novel ligand for HYEM/TR2, stimulates proliferation of T cells and inhibits HT29 cell growth.";

J. Biol. Chem. 273:27548-27556(1998).

[3]

SEQUENCE FROM N.A. (ISOFORM 2), AND PROCESSING.

MEDLINE=21528948; PubMed=11673523;

Granger S.W., Butrovich K.D., Houshmand P., Edwards W.R., Ware C.F.;

"Genomic characterization of LIGHT reveals linkage to an immune response locus on chromosome 19p13.3 and distinct isoforms generated by alternate splicing or proteolysis.";

J. Immunol. 167:5122-5128(2001).

[4]

SEQUENCE FROM N.A.

TISSUE=Brain:

Strausberg R.;

Submitted (Dec-2001) to the EMBL/GenBank/DBJ databases.

-1- FUNCTION: Cytokine that binds to TNFRSF3/LTR. Binding to the decoy receptor TNFRSF6B modulates its effects. Activates NFkB, stimulates the proliferation of T cells, and inhibits growth of the adenocarcinoma HT-29. Acts as a receptor for Herpes simplex virus.

-1- SUBUNIT: Homotrimer.

-1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (isoform 1); Cytoplasmic (isoform 2).

-1- ALTERNATIVE PRODUCTS: 2 isoforms: 1 (shown here) and 2/LIGHT delta-TM; are produced by alternative splicing.

-1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE SPLEEN BUT ALSO TISSUES AND IN HEART, PLACENTA, LIVER, LUNG, APPENDIX, AND KIDNEY, AND NO EXPRESSION SEEN IN FETAL TISSUES, ENDOCRINE GLANDS, OR NONHEMATOPOIETIC TUMOR LINES.

-1- INDUCTION: UPREGULATED AFTER T-CELL ACTIVATION.

-1- PFM: N-glycosylated.

-1- PFM: The soluble form of isoform 1 derives from the membrane form by proteolytic processing.

-1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

-1- CAUTION: Ref.4 sequence differs from that shown due to a frameshift in position 178.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL: AF036581; AAC39563.1; -

EMBL: AF064090; AAC25169.1; -

EMBL: AY028261; AAK26160.1; -

EMBL: BC018058; AAH18058.1; ALT\_FRAME.

HSSP: P01375; 4TSV.

DR Genew: HGNC:11930; TNFSF14.

MM: 604520; -

DR InterPro: IPR003636; TNF\_abcd.  
 DR InterPro: IPR000478; TNF\_family.  
 DR Pfam: PF00229; TNF\_1.  
 DR ProDom: PD002012; TNF\_abcd; 1.  
 DR SMART: SM00207; TNF\_1.  
 DR PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
 DR PROSITE: PS0049; TNF\_2; 1.  
 DR Cytokine: Cytokine; Transmembrane; glycoprotein; signal-anchor;  
 KW Alternative splicing.  
 FT CHAIN 1 240  
 FT  
 FT CHAIN 783 240  
 FT  
 FT DOMAIN 1 37  
 FT TRANSMEM 38 58  
 FT  
 FT DOMAIN 59 240  
 FT SITE 82 83  
 FT DISULFID 154 187  
 FT CARBOHYD 102 102  
 FT VARSPLIC 38 73  
 FT CONFLICT 120 120  
 FT CONFLICT 214 214  
 FT SEQUENCE 240 AA: 26351 MW: 490067ELI390B39 CRC64;

Query Match 7.48; Score 86.5; DB 1; Length 240;  
 Best Local Similarity 22.28; Pred. No. 1.2;  
 Matches 57; Conservative 30; Mismatches 87; Indels 83; Gaps 10;

QY 3 SLGLALACIGLLVYSLG-----SMATLSAQEPSEQLTADNDRREPRLMPQTEESDDV 58  
 DB 32 SCVAVRGDLGLLLMLAGLAVGQWFLQJLHMLGEMVT-----RLP-----DGA 77  
 QY 59 PLEQLVPRRPAKGRKAPRRRAAHYVHPRGDGAQAGVDTGVSEETKINSS 118  
 DB 78 GSWEQLIQERSS-----HEVNPAAHLTGNSSLTG-----SGG 110  
 QY 119 PLKRYDQI-----GEFTVIRAGLYLYCOVHEPE-----GKAYTL-----153  
 DB 111 PLMEQQLGIAFRLGSLYHNDGALVVRKAGYYYSKVQLGVGCPGLASTITRGLYKRT 170  
 QY 154 -----KLDLLNGVIALRCLEESATNASSPQRLRCOVGSLPLRPSSLRITLPA 208  
 DB 171 PRYPEBELLYS-----QSPGCRATSSSRWMDSSFLGYYHLAGEEYVYRVDER 223  
 QY 209 HLKAP-FLTYEGLFOY 224  
 DB 224 LVRLRDGTRSYGAFVW 240

RESULT 12  
 B3A2\_MOUSE  
 ID B3A2\_MOUSE STANDARD; PRT; 1237 AA.  
 AC P13808; Q9ESI3; Q9ESI2; Q9ESI1; Q9ES10; Q9ES09;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Anion exchange protein 2 (Non-erythrocyte band 3-like protein) (B3AP).  
 OS SLCA42 OR AE2.  
 GN Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM A).  
 RX MEDLINE=89034212; PubMed=3182834;  
 RA Alper S.L., Kopito R.R., Lipresco S.M., Lodish H.F.;  
 RT Cloning and characterization of a murine band 3-related cDNA from  
 RT kidney and from a lymphoid cell line.\*;  
 RL J. Biol. Chem. 263:17092-17099(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.  
 RX MEDLINE=20462926; PubMed=11006093;

RA Iecanda J., Urtasun R., Medina J.F.;  
 RT "Molecular cloning and genomic organization of the mouse AE2 anion  
 RT exchanger gene.";  
 RL Biochem. Biophys. Res. Commun. 276:117-124(2000).  
 CC  
 CC -1- FUNCTION: PLASMA MEMBRANE ANION EXCHANGE PROTEIN OF WIDE  
 CC DISTRIBUTION.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS: 5 isoforms; a (shown here), b1, b2, c1 and  
 CC c2; are produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: Isoform a is widely expressed at similar  
 CC levels in all tissues examined. Isoforms b1 and b2 are  
 CC predominantly expressed in stomach although they are also detected  
 CC at lower levels in other tissues. Isoform c1 is stomach-specific.  
 CC Isoform c2 is expressed at slightly higher levels in lung and  
 CC stomach than in other tissues.  
 CC -1- SIMILARITY: BELONGS TO THE ANION EXCHANGER FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: J04036; AAA5505.1; -  
 DR EMBL: AF255774; AAG23154.1; -  
 DR EMBL: AF255774; AAG23155.1; -  
 DR EMBL: AF255774; AAG23156.1; -  
 DR EMBL: AF255774; AAG23158.1; -  
 DR EMBL: AF255774; AAG23157.1; -  
 DR PIR: A31789; A31789.  
 DR HSSP: P02730; 1B7Q.  
 DR -MGD: MGI:109351; SLCA42.  
 DR InterPro: IPR001717; Anion exchange.  
 DR InterPro: IPR003020; HCO3\_cotransp.  
 DR Pfam: PF00955; HCO3\_cotransp. 5.  
 DR PRINTS: PR01231; HCO3TRANSPORT.  
 DR TIGRfam: TIGR00834; ae; 1.  
 DR PROSITE: PS00219; ANION\_EXCHANGER\_1; 1.  
 DR PROSITE: PS00220; ANION\_EXCHANGER\_2; 1.  
 DR Transmembrane: Glycoprotein; Transport; Antipori; Ion transport;  
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.  
 FT DOMAIN 1 703  
 FT  
 FT DOMAIN 704 1237  
 FT TRANSMEM 704 727  
 FT TRANSMEM 733 770  
 FT TRANSMEM 790 812  
 FT TRANSMEM 822 843  
 FT TRANSMEM 844 896  
 FT TRANSMEM 897 914  
 FT TRANSMEM 915 929  
 FT TRANSMEM 930 950  
 FT TRANSMEM 984 1006  
 FT TRANSMEM 1032 1053  
 FT TRANSMEM 1087 1132  
 FT TRANSMEM 1159 1195  
 FT DOMAIN 5 316  
 FT DOMAIN 73 87  
 FT DOMAIN 85 85  
 FT CARBOHYD 855 855  
 FT CARBOHYD 866 866  
 FT CARBOHYD 878 878  
 FT LIPID 1169 1169  
 FT VARSPLIC 1 17  
 FT  
 FT VARSPLIC 1 17  
 FT VARSPLIC 1 166  
 FT VARSPLIC 1 196  
 FT VARSPLIC 1 193  
 FT CONFLICT 205 205  
 FT SEQUENCE 1237 AA: 136813 MW: 1A0782C0071782EE CRC64;



RA Gattl R.A., Derauf D.C., Belmont J.W., Conley M.E.;  
 RT "A single strand conformation polymorphism study of CD40 ligand.  
 RT Efficient mutation analysis and carrier detection for X-linked hyper  
 RT IGM syndrome.";  
 RL J. Clin. Invest. 97:196-201(1996).  
 RP [15]  
 RP VARIANTS HIGM1 ARG-36; CYS-140; SER-231; MET-254 AND GLY-227 DEL.  
 RX MEDLINE-97295077; PubMed-9150729;  
 RA Nonoyama S., Shimadzu M., Toru H., Seyama K., Nuno H., Neubauer M.,  
 RA Yata J.-I., Ochi H.D.;  
 RT "Mutations of the CD40 ligand gene in 13 Japanese patients with  
 RT X-linked hyper-igm syndrome.";  
 RL Hum. Genet. 99:624-627(1997).  
 CC -1- FUNCTION: MEDIATES B-CELL PROLIFERATION IN THE ABSENCE OF CO-  
 CC STIMULUS AS WELL AS IGE PRODUCTION IN THE PRESENCE OF IL-4.  
 CC INVOLVED IN IMMUNOGLOBULIN CLASS SWITCHING.  
 CC -1- SUBUNIT: HOMOTRIMER.  
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN  
 CC EXTRACELLULAR SOLUBLE FORM.  
 CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED ON ACTIVATED CD4+  
 CC T-LYMPHOCYTES.  
 CC -1- PTM: The soluble form derives from the membrane form by  
 CC proteolytic processing.  
 CC -1- DISEASE DEFECTS IN TNFSF5 ARE THE CAUSE OF AN X-LINKED  
 CC IMMUNODEFICIENCY WITH HYPER-IGM (HIGM1), AN IMMUNOGLOBULIN ISOTYPE  
 CC SWITCH DEFECT CHARACTERIZED BY ELEVATED CONCENTRATIONS OF SERUM  
 CC IGM AND DECREASED AMOUNTS OF ALL OTHER ISOTYPES. AFFECTED MALES  
 CC PRESENT AT AN EARLY AGE (USUALLY WITHIN THE FIRST YEAR OF LIFE)  
 CC RECURRENT BACTERIAL AND OPORUNISTIC INFECTIONS, INCLUDING  
 CC PNEUMOCYSTIS CARINI PNEUMONIA AND INTRACTABLE DIARRHEA DUE TO  
 CC CRYPTOSPORIDIUM INFECTION. DESPITE SUBSTITUTION TREATMENT WITH  
 CC INTRAVENOUS IMMUNOGLOBULIN, THE OVERALL PROGNOSIS IS RATHER POOR,  
 CC WITH A DEATH RATE OF ABOUT 10% BEFORE ADULESCENCE.  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -1- DATABASE: NAME-CD40Lbase;  
 CC NOTE-European CD40L defect database (mutation db);  
 CC WWW="http://www.expasy.org/databases/cd40lbase/";  
 CC FTP="ftp://ftp.expasy.org/databases/cd40lbase/";  
 CC -1- DATABASE: NAME-PROW; NOTE-CD guide CD154 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd154.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC -----  
 DR EMBL: X68550; CAA48554.1; -  
 DR EMBL: Z15017; CAA48737.1; -  
 DR EMBL: X67818; CAA48077.1; -  
 DR EMBL: L07414; AAA35662.1; -  
 DR EMBL: D31797; BAA06599.1; -  
 DR EMBL: D31793; BAA06599.1; JOINED.  
 DR EMBL: D31794; BAA06599.1; JOINED.  
 DR EMBL: D31795; BAA06599.1; JOINED.  
 DR EMBL: D31796; BAA06599.1; JOINED.  
 DR PIR: S25684; S25684.  
 DR PIR: S25694; S25694.  
 DR PIR: S28017; S28017.  
 DR PIR: S28852; S28852.  
 DR PIR: JH0793; JH0793.  
 DR PDB: 1ALY; 17-SEP-97.  
 DR GeneW: HGNC:11935; TNFSF5.  
 DR MIM: 300386; -  
 DR MIM: 308230; -  
 DR InterPro: IPR003263; TNF\_5.  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR004478; TNF\_family.  
 DR Pfam: PF00229; TNF; 1.  
 DR ProDom: PD008600; TNF\_5; 1.  
 DR SMART: SM00207; TNF; 1.

DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS0049; TNF\_2; 1.  
 DR Cytokine; Transmembrane; Glycoprotein; Signal-anchor; Antigen;  
 KW disease mutation; Polymorphism; 3d-structure.  
 FT CHAIN 1 261 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 Query Match 7.4%; Score 86; DB 1; Length 261;  
 Best Local Similarity 25.9%; Pred. No. 1.4;  
 Matches 35; Conservative 21; Mismatches 55; Indels 24; Gaps 6;  
 QY 105 TVSGWE---TKINSSPLRYRQIGETVTRAGLYLYCQVHPDEGRAVYIKDLLVN 160  
 DB 136 SYLQWAEKGYTWSNNLVLENGKQ--TVKQGLYLYAYQVFCNSRNASQAPFIAS 192  
 QY 161 GVALACLEEF-----SATASPPQPLRIC-----QVSGLLPLRPGSSLRIRTPMAH 209  
 DB 193 --LCIKSPGFERILLRANTHSARP-----CGQOSIHGVGEVLPASVFNATDPSQ 246  
 QY 210 LKAAPLYTFGFQV 224  
 DB 247 VSHGTGFTSGFLKL 261  
 RESULT 14  
 ID TNFS\_MACMU STANDARD; PRT; 261 AA.  
 AC 09BOC7;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-  
 DE L) (CD154 protein).  
 GN TNFSF5 OR CD40LG OR CD40L.  
 OS Macaca mulatta (Rhesus macaque), and  
 OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OC NCBI\_TaxID=9544, 9531;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND M.MULATTA VARIANTS GLN-60; VAL-204; PRO-206  
 RP AND THR-215.  
 RC SPECIES-M.mulatta, and C.torquatus atys; TISSUE-Lymphocytes;  
 RX MEDLINE-21583618; PubMed-11491535;  
 RA Villinger F., Bostlik P., Mayne A.E., King C.L., Genain C.P.,  
 RA Weiss W.R., Ansari A.A.;  
 RT "Cloning, sequencing and homology analysis of nonhuman primate  
 RT Fas/FasL ligand and co-stimulatory molecules.";  
 RL Immunogenetics 53:315-328(2001).  
 CC -1- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell  
 CC proliferation in the presence of co-stimulus as well as IGE  
 CC production in the presence of IL-4. Involved in immunoglobulin  
 CC class switching (By similarity).  
 CC -1- SUBUNIT: Homotrimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an  
 CC extracellular soluble form (By similarity).  
 CC -1- PTM: The soluble form derives from the membrane form by  
 CC proteolytic processing (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC -----  
 DR EMBL: AF344859; AAK37541.1; -  
 DR EMBL: AF344841; AAK37600.1; -  
 DR HSSP: P29965; 1ALY.  
 DR InterPro: IPR003263; TNF\_5.  
 DR InterPro: IPR003636; TNF\_abc.

DR InterPro: IPR00478; TNF family.  
 DR Pfam: PF00229; TNF\_1.  
 DR Prodom: PD002012; TNF\_abc; 1.  
 DR Prodom: PD008600; TNF\_5; 1.  
 DR SMART: SM00207; TNF\_1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 DR Cytokine; Transmembrane; Glycoprotein; Signal-anchor; Polymorphism.  
 FT CHAIN 1 261  
 FT CHAIN 113 261  
 FT DOMAIN 1 22  
 FT TRANSMEM 23 43  
 FT SITE 44 261  
 FT DISULFID 112 113  
 FT CARBOHYD 178 218  
 FT VARIANT 240 240  
 FT VARIANT 60 60  
 FT VARIANT 204 204  
 FT VARIANT 206 206  
 FT VARIANT 215 215  
 SQ SEQUENCE 261 AA; 29366 MW; AA2E0F11C0F28ABB CRC64;  
 Query Match  
 Best Local Similarity 7.4%; Score 86; DB 1; Length 261;  
 Matches 35; Conservative 21; Mismatches 55; Indels 24; Gaps 6;  
 OY 105 TVSGWEE---TKINSSPLRYDROIGFTVIRAGLYLYCOVHFDEGRAVYLKLDLVN 160  
 DB 136 SVLQNAERKGYTMSNNLVLENGKQL---TVKROGLYIYAQVFCSNREASSQAPFIAS 192  
 OY 161 GYLARCLEEF-----SATASSPGPOLRLC-----QVSGLLPLRGSSLRIRTPMAH 209  
 DB 193 --LCIKNSPGRFERILLRAANTHSSAKP---CGQOSIHIGVFELOPGASVFWVTDPDPO 246  
 OY 210 LKAAPLYFGFLFOV 224  
 DB 247 VSHGTCFTSFGLKL 261  
 RESULT 15  
 TNFS\_MACNE  
 ID TNFS\_MACNE STANDARD; PRT; 240 AA.  
 AC 09BDM7;  
 DT 15-JUN-2002 (rel. 41, Created)  
 DT 15-JUN-2002 (rel. 41, Last sequence update)  
 DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-  
 l) (CD154 protein).  
 GN TNFSF5 OR CD40LG OR CD40L.  
 OS Macaca nemestrina (Pig-tailed macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9545;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lymphocytes;  
 RX MEDLINE=21383618; PubMed=11491535;  
 RA Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,  
 RA Weiss W.R., Ansari A.A.;  
 RT "Cloning, sequencing, and homology analysis of nonhuman primate  
 RT Fas/Fas-ligand and co-stimulatory molecules.";  
 RL Immunogenetics 53:315-328(2001).  
 CC -1- FUNCTION: Cytokine that binds to TNFSF5. Mediates B-cell  
 CC proliferation in the absence of co-stimulus as well as Ige  
 CC production in the presence of IL-4. Involved in immunoglobulin  
 CC class switching (By similarity).  
 CC -1- SUBUNIT: Homotrimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an  
 CC extracellular soluble form (By similarity).

CC -1- PTM: The soluble form derives from the membrane form by  
 CC proteolytic processing (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).  
 CC -----  
 DR EMBL: AF344853; AAK37613.1; -  
 DR HSSP: P29965; 1ALY.  
 DR InterPro: IPR003263; TNF\_5.  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR00478; TNF\_family.  
 DR Pfam: PF00229; TNF\_1.  
 DR Prodom: PD002012; TNF\_abc; 1.  
 DR Prodom: PD008600; TNF\_5; 1.  
 DR SMART: SM00207; TNF\_1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 DR Cytokine; Transmembrane; Glycoprotein; Signal-anchor.  
 FT CHAIN 1 240  
 FT CHAIN 113 240  
 FT DOMAIN 1 22  
 FT TRANSMEM 23 43  
 FT SITE 44 240  
 FT DISULFID 112 113  
 FT CARBOHYD 173 197  
 FT VARIANT 219 219  
 SQ SEQUENCE 240 AA; 27273 MW; BA6FA0213CAEDC CRC64;  
 Query Match  
 Best Local Similarity 7.3%; Score 85; DB 1; Length 240;  
 Matches 34; Conservative 20; Mismatches 53; Indels 24; Gaps 6;  
 OY 109 WEB---TKINSSPLRYDROIGFTVIRAGLYLYCOVHFDEGRAVYLKLDLVN 164  
 DB 119 IAKGYYTMSNNLVLENGKQL---TVKROGLYIYAQVFCSNREASSQAPFIAS--LC 173  
 OY 165 LKCLEEF-----SATASSPGPOLRLC-----QVSGLLPLRGSSLRIRTPMAH 213  
 DB 174 LKSPGRFERILLRAANTHSSKTP---CGQOSIHIGVFELOPGASVFWVTDPDPOVSHG 229  
 OY 214 PFLTYFGFLFOV 224  
 DB 230 TGFTSFGLKL 240

Search completed: May 8, 2003, 02:20:14  
 Job time : 10.3988 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:13:22 ; Search time 12.3772 Seconds

(without alignments)  
1747.586 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162

Sequence: 1 VLSIGLALACILGLLVVSL.....PMAHLKAPLFTYFGLEQVH 225

Scoring table:

BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90.5	7.8	210	2 D87394	hypothetical prote
2	90.5	7.8	421	2 C83040	hypothetical prote
3	89	7.7	352	2 T08209	matrix protein - H
4	88	7.6	532	2 S12076	alkaline phosphata
5	87.5	7.5	670	2 A87348	alpha-glucosidase,
6	87	7.5	244	2 A46066	lymphotoxin beta -
7	86.5	7.4	340	2 A83035	probable transmem
8	86.5	7.4	1237	2 A31789	band 3-related pro
9	86	7.4	261	2 A53476	band 3-related - huma
10	85	7.3	1237	2 A56764	band 3-related pro
11	84	7.2	278	2 A49266	fas ligand - rat
12	84	7.2	2962	2 T19756	hypothetical prote
13	83.5	7.2	394	2 H95947	probable conserved
14	83	7.1	467	2 T34874	hypothetical prote
15	82	7.1	335	1 PAHUA	alkaline phosphata
16	82	7.1	654	2 S76870	hypothetical prote
17	80.5	6.9	246	2 S29328	complement subcomp
18	80.5	6.9	1234	2 A34911	band 3-related pro
19	80.5	6.9	1438	2 T17402	band 3-related pro
20	80	6.9	261	2 AB3070	conserved hypotet
21	80	6.9	319	2 H98216	hypothetical prote
22	79.5	6.8	197	2 JH0309	tumor necrosis fac
23	79	6.8	487	2 F82065	RNA polymerase sig
24	79	6.8	944	2 S75188	DNA topoisomerase
25	79	6.8	1584	2 JN0114	kinesin-related pr
26	79	6.8	1584	2 JN0114	kinesin-like prote
27	78.5	6.8	289	2 H83433	hypothetical prote
28	78.5	6.8	774	2 D83507	hypothetical prote
29	78.5	6.8	1164	2 B71429	phytochrome D - Ar

30	78.5	6.8	1396	2 A44453	translaton initia
31	78	6.7	240	2 T49856	autoyhsn, N-acety
32	78	6.7	590	2 AB1411	amyloid beta (A4)
33	77.5	6.7	763	2 A49321	nuclear autoantige
34	77.5	6.7	787	2 A48819	Fas ligand - mouse
35	77	6.6	279	2 A53062	smg protein - Esch
36	76.5	6.6	214	2 A26227	probable lipote-p
37	76.5	6.6	562	2 A98297	nuclear antigen EB
38	76.5	6.6	562	2 C86138	hypothetical prote
39	76.5	6.6	1069	2 S27922	hypothetical prote
40	76	6.5	200	2 B83642	beta-amyloid precu
41	76	6.5	380	2 T34568	hypothetical prote
42	76	6.5	511	2 J01404	CDER-box DNA-bindi
43	76	6.5	751	2 A49274	hypothetical prote
44	75.5	6.5	466	2 T24257	hypothetical prote
45	75.5	6.5	575	1 WFBOM	mullerian inhibiti

#### ALIGNMENTS

##### RESULT 1

D87394 hypothetical protein CC1168 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C:Accession: D87394

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg

B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; K

n., J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser,

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249, MUID:21173696, PMID:11259647

A:Accession: D87394

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-210 <STO>

A:Cross-references: GB:AE005673; NID:G13422488; PIDN:ANK23152.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC1168

Query Match 7.8%; Score 90.5; DB 2; Length 210;  
Best Local Similarity 26.2%; Pred. No. 1.6;  
Matches 33; Conservative 19; Mismatches 67; Indels 7; Gaps 2;

QY	2	LSGLIALACILGLLVVSLGSWATLSAQEPSEELTAEDRREPPPLNPQTESQDVYFL	61
DB	32	LRIGLAVFAAGVAAATVIVQAAWRSLSKLTQATATLVLDKPRF--TGVLKDGRRPFL	88
QY	62	EQLVPRRSAPKGRKARRRAIAAHYEVHPRPGDGAQAGV---DGTVSGMEETKINSS	117
DB	89	ITAEAEEDAKDNITVRLTAPLVLRGYPNPSQATKSGVYREANTLLLTDEVKITSA	148
QY	118	SPLRVD 123	
DB	149	EGDFPD 154	

##### RESULT 2

C83040 hypothetical protein PA4851 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000

C:Accession: C83040

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.;

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic p

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: C83040

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-421 <STO>  
 A:Cross-references: GB:AE004898; GB:AE004091; NID:g9951115; PIDN:AA08236.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA4851

Query Match 7.8%; Score 90.5; DB 2; Length 421;  
 Best Local Similarity 22.6%; Pred. No. 3.4;  
 Matches 53; Conservative 27; Mismatches 76; Indels 79; Gaps 10;

29 QESQSELTAEADR-----RPPPLNPTQTESQOVVFFL-----EQLV 65  
 DB 163 QESQSELTAEADR-----RPPPLNPTQTESQOVVFFL-----EQLV 65  
 QY 66 RPRRAAFKGRKARRRAIAHVEHPRGDGAQAGVGTSGWEETKINSSPLRYDRQ 125  
 DB 223 EDSVAPALGSAKRPDEA-----PVESLQDLNDEPIQLAMEKR-----RQMPRR 266  
 QY 126 IGEFTYV---RAGLYTYICQVHDE-----GKAVYTKLDL-LVNGV 162  
 DB 267 LGMULLLILALGILAAQYIAYHDELLARODAYRPWFQALCPETIGCTLPKRYDVEQIRSSN 326  
 QY 163 LALRCLSEFSA-----TAASSGP--QLRLQVSGIL---PLRPGSSL 200  
 DB 327 LVVRSHEFSGALVVDIAITNKRASFQPFPLLELRFADLNGHLIARRRFRGETL 381

RESULT 3  
 T08209  
 matrix protein - Hendra virus  
 C:Species: Hendra virus  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
 C:Accession: T08209  
 R:Yu, M.; Hanson, E.; Langedijk, J.P.M.; Eaton, B.T.; Wang, L.F.  
 Virology 251, 227-233, 1998  
 A:Title: The attachment protein of hendra virus has high structural similarity but limited  
 A:Reference number: Z16405; MID:99058172; PMID:9637786  
 A:Accession: T08209  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: genomic RNA  
 A:Residues: 1-352 <YUM>  
 A:Cross-references: EMBL:AF017149; NID:g3273489; PIDN:AA083191.1; PID:g3273494  
 C:Genetics:  
 A:Gene: M  
 C:Superfamily: parainfluenza virus matrix protein  
 C:Keywords: matrix protein

Query Match 7.7%; Score 89; DB 2; Length 352;  
 Best Local Similarity 28.1%; Pred. No. 3.8;  
 Matches 39; Conservative 20; Mismatches 50; Indels 30; Gaps 8;  
 QY 78 RPRRAIAHVEHPRGDGAQAGVGTSGWEETKINSSPLRYDRQIGFTYIRAGLY 137  
 DB 197 RNNNAIAFNLLYTKIDADLAKAGIGSFDK-DGTVAS-----FMHLGNF-VRRGKY 249  
 QY 138 YL--YCVHDEGKAVY-----LKIDLVNGVLAIRCLSEFSAPRASSPQQLRQV 188  
 DB 250 YSEYECRRKIDRMKQLFSLGISGLSLIKINGVISKRLFAQM-----GLQKNLC-- 299  
 QY 189 SGLPLRPGSSLRIPTLPW 207  
 DB 300 FSLMDINFM---LNRLTW 314

RESULT 4  
 S12076  
 alkaline phosphatase (EC 3.1.3.1) precursor, placental-like - human  
 N:Alternate names: alkaline phosphatase Nagao isozyme; germ-cell alkaline phosphatase  
 C:Species: Homo sapiens (man)  
 C:Date: 21-Nov-1993 #sequence\_revision 13-Mar-1998 #text\_change 31-Jan-2000  
 C:Accession: S12076; 137991; A34333; S00974; A31336  
 R:Gum, J.R.; Hicks, J.W.; Sack, T.L.; Kim, Y.S.  
 Cancer Res. 50, 1085-1091, 1990

A:Title: Molecular cloning of complementary DNAs encoding alkaline phosphatase in hum

A:Reference number: S12076; MID:90124311; PMID:2297757

A:Accession: S12076

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-177, 'A', 179-272, 'W', 274-497, 'S', 499-530, 'T', 532 <GUM>

A:Cross-references: EMBL:X53279; NID:g28624; PIDN:CAA3734.1; PID:g28625

R:Lower, M.E.; Straus, A.W.

Cancer Res. 50, 3956-3962, 1990

A:Title: Expression of a Nagao-type, phosphatidylinositol-glycan anchored alkaline ph

A:Reference number: 137991; MID:90283879; PMID:2162249

A:Accession: 137991

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-151, 'V', 153-532 <LOW>

A:Cross-references: EMBL:X55958; NID:g296797; PIDN:CAA39425.1; PID:g296798

R:Matnabe, S.; Matnabe, T.; Li, W.B.; Soong, B.W.; Chou, J.Y.

J. Biol. Chem. 264, 12611-12619, 1989

A:Title: Expression of the germ cell alkaline phosphatase gene in human chorioncarcino

A:Reference number: A34333; MID:89308696; PMID:2745460

A:Accession: A34333

A:Molecule type: mRNA

A:Residues: 1-379, 'L', 381-497, 'R', 499-532 <MAT>

A:Cross-references: GB:J04948; NID:9178418

R:Shen, L.P.; Liu, H.; Kan, Y.W.; Kam, W.

Nucleic Acids Res. 16, 5694, 1988

A:Title: 5' nucleotide sequence of a putative human placental alkaline phosphatase-11

A:Reference number: S00974; MID:88262578; PMID:3387245

A:Accession: S00974

A:Molecule type: DNA

A:Residues: 1-157 <SHE>

A:Cross-references: EMBL:X07247; NID:g35509

A:Note: The authors translated the codon CCG for residue 28 as Cys and AGT for residu

R:Millan, J.L.; Manes, T.

Proc. Natl. Acad. Sci. U.S.A. 85, 3024-3028, 1988

A:Title: Seminoma-derived Nagao isozyme is encoded by a germ-cell alkaline phosphatas

A:Reference number: A31336; MID:88203632; PMID:2834730

A:Accession: A31336

A:Molecule type: DNA

A:Residues: 1-56, 'W', 58-151, 'V', 153-315, 'L', 317-532 <MIT>

A:Cross-references: GB:J03252; NID:g178427; PIDN:AAA98616.1; PID:g178428; GB:J03512

C:Genetics: 23/1; 62/1; 100/3; 159/1; 216/3; 261/3; 286/1; 331/1; 395/1; 431/1

A:Note: the first intron occurs before the initiator codon

C:Function:

A:Description: catalyzes the hydrolysis of phosphate monoesters

C:Superfamily: alkaline phosphatase

C:Keywords: blocked carboxyl end; glycoprotein; homodimer; lipoprotein; membrane prot

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-503/Product: alkaline phosphatase, placental-like #status predicted <MAT>

F:504-532/Domain: carboxyl-terminal propeptide #status predicted <CTP>

F:111/Active site: Ser (phosphoserine intermediate) #status predicted

F:141,268/Binding site: carboxylate (Asn) (covalent) #status predicted

F:185/Binding site: substrate phosphate (Arg) #status predicted

F:339,379,451/Binding site: zinc (His) #status predicted

F:503/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Asp) (in mature f

Query Match 7.6%; Score 88; DB 2; Length 532;  
 Best Local Similarity 28.6%; Pred. No. 7.3;  
 Matches 44; Conservative 16; Mismatches 44; Indels 50; Gaps 8;

QY 76 KARRRAIAHVEV---HPRPG---ODGAQAGVGTSGWEETKINSSPLRYDRQIGF 129  
 DB 400 KARDRA---YTVLLYNGPGVGLDGAAPDVTESSGSPVYRQOSAVPLDDETHAGD 455  
 QY 130 TVRAGLYTYICQVHDEGKAVYTKIDLAVNGV-----LALRCLSEFSA----- 173  
 DB 456 VAV-----FARGQAH-----LVHGQDQFTAHVMAFAACLEPYACDLAP 498  
 QY 174 ---TAASSPQQLRQVSGILPLRPGSSLRIPT 204  
 DB 499 AGTTDAHHPGV---VPALPLLAGTLLLTGT 528



## RESULT 5

A87348

alpha-glucosidase, probable [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C:Accession: A87348

R.Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.J.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: A87348

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-670 &lt;STO&gt;

A:Cross-references: GB:AE005673; NID:g13422039; PIDN:AAK22781.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0796

Query Match

Best Local Similarity 7.5%; Score 87.5; DB 2; Length 670;

Matches 43; Conservative 24; Mismatches 75; Indels 37; Gaps 6;

58 VFLEQLVLP-----RRSAPKGRKAR-----PRRAIAHYEHVHRP----- 93

498 LPFTMLGLPMDYTPGGRNVAAPKPKSQFLLPTVQTRGQALMAYVDSPLTIVNSDSP 557

94 -GQDAQAQAVD---GTVSGMEETKINSSPLRYDRQIGFTYI--RAGLYLYCOVHEDE 147

558 IYYAASPAGLDRIASAVPTSMDETRVLSG-----EIGGYIYARRKKADMMVGMATNNA 610

148 GAAVYLAIDLIVNGVIALRCLFEFSATASSPGPOLRCQVSGLLPRLRGSSLRITLP 206

611 GRTVAVPSLFGNAPSAEIRREDGAPETALKTRTQSVASKOTPLTKLAPGGGVIRISP 669

Db

## RESULT 6

A46066

Lymphotoxin beta - human

C:Species: Homo sapiens (man)

C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000

C:Accession: A46066

R.Browning, J.L.; Ngam-ek, A.; Lawton, P.; Demarinis, J.; Tilzard, R.; Chow, E.P.; Hessio

Cell 72, 847-856, 1993

A:Title: Lymphotoxin beta, a novel member of the TNF family that forms a heteromeric com

A:Reference number: A46066; MUID:93208881; PMID:7916655

A:Accession: A46066

A:Status: preliminary

A:Molecule type: DNA; protein

A:Residues: 1-244 &lt;BRO&gt;

A:Cross-references: GB:L11015; NID:g2922276; PIDN:AAA36191.1; PID:g2922277

A:Note: sequence extracted from NCBI backbone (NCBIN:128066, NCBIPI:128067)

C:Keywords: transmembrane protein

Query Match

Best Local Similarity 7.5%; Score 87; DB 2; Length 244;

Matches 64; Conservative 26; Mismatches 80; Indels 108; Gaps 13;

3 SIGLALA---CLGLLLVNVSLGSMATLSAQSPQSELTAEADR-----REPEPL 47

18 SLLAVAGTSLVTLTLLAVPIVLAIVLPDQGGIVTETDPGAGQAGGIFGPKLPBE 77

48 NPQTESQDVVPFLPOLVPRPSAPKGRKARRRAIAHYEHVPRPGDGAQAGVDGTVS 107

78 EPEET---DLSPGLP-----AAHLIGAPLKGGQ-----L 102

Db

## RESULT 7

A83035

probable transmembrane sensor PA4895 [imported] - Pseudomonas aeruginosa (strain PAC

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 15-Jun-2001

C:Accession: A83035

R.Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.;

Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic p

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: A83035

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-340 &lt;STO&gt;

A:Cross-references: GB:AE004902; GB:AE004091; NID:g9951162; PIDN:AG08280.1; GSPDB:G

C:Genetics:

A:Gene: PA4895

C:Superfamily: Pseudomonas putida regulatory protein pupr

Query Match

Best Local Similarity 7.4%; Score 86.5; DB 2; Length 340;

Matches 67; Conservative 26; Mismatches 85; Indels 93; Gaps 14;

5 GLAACLGLLV-----VSLGSMATLSAQSPQSEEL----- 36

104 GLALLGLAPLVWGMGMAQVWRDGFCAVDLTAVERRDVLDEDSQVEENTDSALDVRD 163

37 -----TAEDEPEELNPQTESQDVVPFL-----EQLYRPRKAPKGRK 76

164 AGORILLRYRGEIYRTADREPR-----PELVTRDQGLRALGTFVSRR 211

77 APRRAIAHYE--VHPRPGDGAQAGVDGTVSGMEETKINSSPLRYDRQIGFTYIRA 134

212 BEAEVIAV-YEGANQVPR--EGASADGRV-----IEGQVRRDRQ-----RI 254

135 GLYLYYCOVHEDEKAVYIKLIDLIVNGVIALRCLFEFSATASSPGPOLRCQVSGLLPL 194

255 G-----PVESASEAALAMROGLLVADDPPLR---QMGELMRYGGSIE-CEPS-LDPL 303

195 RPSSLRITLPMALIKAAPLTITGLFOVH 225

304 RVSGTFPPVDLP---LALAMLAQTHGLRLVH 331

Db

## RESULT 8

A31789

band 3-related protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 21-May-1990 #sequence\_revision 21-May-1990 #text\_change 20-Aug-1999

C:Accession: A31789

R.Alper, S.L.; Kopito, R.R.; Libresco, S.M.; Lodish, H.F.

J. Biol. Chem. 263, 17092-17099, 1988

A:Title: Cloning and characterization of a murine band 3-related cDNA from kidney an

A:Reference number: A31789; MUID:89034212; PMID:3182834

A:Accession: A31789

A:Molecule type: mRNA

A:Residues: 1-1237 &lt;ALP&gt;

A:Cross-references: GB:J04036; NID:g192123; PIDN:AAA65505.1; PID:g309114

C:Superfamily: band 3 anion transport protein

C:Keywords: transmembrane protein

Query Match

Best Local Similarity 7.4%; Score 86.5; DB 2; Length 1237;



Best Local Similarity 22.9%; Pred. No. 7.8;  
Matches 56; Conservative 30; Mismatches 67; Indels 92; Gaps 15;

QY 15 LVVVSIGSNATLSAQEPSEELTAEDRREPELNPOTESQDVVPLEQLVRRNSAPKG 74  
DB 91 LVGMSIGSNATLSAQEPSEELTAEDRREPELNPOTESQDVVPLEQLVRRNSAPKG 136  
QY 75 RARPRRAIAHYEVHPRRGQDGAQAGVGTGSGMEETKINS-SSPLARDROIGETVTR 133  
DB 137 TK-KPRS--VAHLTGMPKRSRSTPLE-----MEDYGTALISGVTKK--GGIIVNE 182  
QY 134 AGLVLYLCOVHDEKRA-----VYLKLDLVNGVALRCLEEFSAATASSPQQLR 184  
DB 183 AGLVLYLCOVHDEKRA-----VYLKLDLVNGVALRCLEEFSAATASSPQQLR 228  
QY 185 LCQVSGFLPLRPGSSLRITLPMWMA-----LKAPPL-----TYF 219  
DB 229 YCTTGQI-----WAHSSYLGAVFNLTVADHLVYNISQSLINEESKTFE 273  
QY 220 GLEOV 224  
DB 274 GLTKL 278

## RESULT 12

hypothetical protein C35C5.1 - *Caenorhabditis elegans*

C.Species: *Caenorhabditis elegans*  
C.Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000

C.Accession: T19756; T21561  
R.White, S.

submitted to the EMBL Data Library, August 1996  
A.Reference number: T19756

A.Accession: T19756  
A.Status: preliminary; translated from GB/EMBL/DBJ

A.Molecule type: DNA  
A.Residues: 1-2962 <W12>

A.Cross-references: EMBL:Z78417; PIDN:CAB01693.1; GSPDB:GN00028; CESP:C35C5.1  
A.Experimental source: clone C35C5

R.Harris, B.  
submitted to the EMBL Data Library, August 1996

A.Reference number: Z19441  
A.Accession: T21561

A.Status: preliminary; translated from GB/EMBL/DBJ  
A.Molecule type: DNA

A.Residues: 1-2962 <W12>

A.Cross-references: EMBL:Z78543; PIDN:CAB01756.1; GSPDB:GN00028; CESP:C35C5.1  
A.Experimental source: clone F2966

C.Genetics:  
A.Gene: CESP:C35C5.1

A.Map position: X  
A.intons: 373/3; 452/2; 525/2; 576/3; 598/3; 782/2; 894/1; 937/3; 1334/3; 1360/3; 1587/

## Query Match

Best Local Similarity 21.6%; Pred. No. 1.1e+02;

Matches 42; Conservative 29; Mismatches 91; Indels 32; Gaps 7;

QY 29 QESQSELTAEADRRREPPELNPOTESQDVV-----PELEQLVRRNSAPK-----GRKA 77  
DB 128 EDLEHLLNAYRHEEEEC---FEEDDITTELPPLVKAQVKKPRKRLPHLSTIESGSTA 184  
QY 78 RRRRAIAHYEVHPRRGQDGAQAGV--DGTSGMEETKI-----SSSPLARDROIGER 129  
DB 185 KSKSLVAEYVHPRVNRKRAVYDDGKVVYQKRTYTRNIMSTIPIQYHLEGETER 244  
QY 130 TVIRAGLYLYCOVHDEKRAVYLKLDLVNGVALRCLEEFSAATASSPQQLRQVNS 189  
DB 245 GRVKESTLSKTEIQYLQAKLVSPKCD-----QFREQIVATAVEYDG--SVKMLQFE 294  
QY 190 GLLPLRPGSSLRIR 203  
DB 295 NALKKSHGKOKRLK 308

## RESULT 13

probable conserved membrane-anchored protein SMD21182 [imported] - *Sinorhizobium meli*

C.Species: *Sinorhizobium meli*  
C.Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

C.Accession: H95947  
R.Finan, T.M.; Weidner, S.; Wong, K.; Buhmester, J.; Chain, P.; Vorholter, F.J.; He

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A.Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing

A.Reference number: A95842; MUID:21396508; PMID:11481431  
A.Accession: H95947

A.Status: preliminary  
A.Molecule type: DNA

A.Residues: 1-394 <KUP>

A.Cross-references: GB:AL591985; PIDN:CAC49248.1; PID:G15140734; GSPDB:GN00167  
A.Experimental source: strain 1021, megaplasmid pSymB

R.Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hub  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R

L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001

A.Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lela  
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yen

A.Title: The composite genome of the legume symbiont *Sinorhizobium meli*lot1.  
A.Reference number: A96039; MUID:21368234; PMID:11474104

A.Contents: annotation  
C.Genetics:

A.Gene: SMD21182  
A.Genome: plasmid

C.Superfamily: Mycobacterium tuberculosis hypothetical protein Rv3272

## Query Match

Best Local Similarity 24.3%; Pred. No. 13;

Matches 63; Conservative 25; Mismatches 92; Indels 79; Gaps 12;

QY 13 LILVVSIGSW-----ATLSAQEPSEELTAEDRRE--PELNPOTESQDVVPLEQLV 65  
DB 11 LELARILAPWIGQITLADGADYIKVESPADGDTFRMGPPVYGEDEERLDAA-YFNACN 69  
QY 66 RRRRS-----APKGRKAPRRRAIA-----HYE-----VHRP----- 93  
DB 70 RGRRSVLDFTTAEGEAVRRLAAQSDVILENFKVGLAKYGLDYSLSKVNRLIYCSV 129  
QY 94 ---GODGA---QAGVGTGSGMEETKINSPLRYDROIG-EFTVIRAGLY-----YL 139  
DB 130 TGRGQDGPYAHRAIGYIYOGMSGIMDLNGEPREOKIKVAPADLFTGLYGVIAQAL 189  
QY 140 YCOVHDEKRAVYLKLDLVNGVALRCLEEFSAATASSPQQLRQVSGLLPLRPGSS 199  
DB 190 AGRERTGEGQIDMALDCMTGVLANQALNFL-----VSGKAPRRIGN- 232  
QY 200 LRIRITLPMWMAIKRAFLTY 218  
DB 233 -----AHPNIAPYOVF 243  
RESULT 14  
T34874  
hypothetical protein SC3A3.09 - *Streptomyces coelicolor*  
C.Species: *Streptomyces coelicolor*  
C.Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
R.Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M  
submitted to the EMBL Data Library, August 1999  
A.Reference number: Z21560  
A.Accession: T34874  
A.Status: preliminary; translated from GB/EMBL/DBJ  
A.Molecule type: DNA  
A.Residues: 1-467 <SEB>  
A.Cross-references: EMBL:AL109849; PIDN:CAB52866.1; GSPDB:GN00070; SCOEDB:SC3A3.09  
C.Genetics:  
A.Experimental source: strain A3(2)  
A.Gene: SCOEDB:SC3A3.09

Query Match 7.1%; Score 83; DB 2; Length 467;  
 Best Local Similarity 23.6%; Pred. No. 17;  
 Matches 57; Conservative 22; Mismatches 91; Indels 72; Gaps 11;

4 LGALACL-----GGLLVVSLGSM---ATLSAQPSQDELTAE--DKREPP-- 45  
 112 LGRVAGCLGKRPVEKPLRPLGRLGGRAGNMPPTAYFARQVPRDLTAHWNRRSAPTS 171  
 46 ---ELNPTESQDVPEPLEQV--RPRRSAPKGRKAR-----PRAIQA----- 85  
 172 LAENIDGPREDDNLNPLGLGILLRRHCKAFTTGDAVAVMDELPRGFTFAERVALNL 231  
 86 -----HYEVHPRPGDGAQAGVGTVSWEETKINSSEPLRDROIGFTVIRAGLYX 138  
 232 LTGLEPPTARHNPFRFEMIGALIRADYHGW---TNGDP-----GAIAE----- 273  
 139 LYCOVHPEBGKAVYIKLDLVNGLALRCLKEFSPTAASSFGPQLRCLCVGLPLRPGS 198  
 274 ---QAHRD-----ATLSHTANGVAHMAFMAAVALVAAAPDDAPDVHACLRAGLAVVPPGS 324

OY 199 SL 200  
 Db 325 RL 326

RESULT 15  
 PAMHA  
 alkaline phosphatase (EC 3.1.3.1) precursor, placental (validated) - human  
 N:Alternate names: glycerophosphatase; orthophosphoric-monoester phosphohydrolase, Regan  
 C:Species: Homo sapiens (man)  
 C:Date: 30-Sep-1987 #sequence, revision 28-Jul-1995 #text, change 08-Dec-2000  
 C:Accession: A31074; A27363; A24318; B24318; A25385; A94059; A28958; A90114; A24  
 R:Knoll, B.J.; Rothblum, K.N.; Longley, M.  
 J. Biol. Chem. 263, 12020-12027, 1988  
 A:Title: Nucleotide sequence of the human placental alkaline phosphatase gene. Evolution  
 A:Reference number: A31074; MUID:8829886; PMID:3042787  
 A:Accession: A31074  
 A:Molecule type: DNA  
 A:Residues: 1-535 <KNO>  
 A:Cross-references: GB:J03931; GB:M19159; NID:g178475; PIDN:AAA51710.1; PID:g178476  
 R:Knoll, B.J.; Rothblum, K.N.; Longley, M.  
 Gene 60, 267-276, 1987  
 A:Title: Two gene duplication events in the evolution of the human heat-stable alkaline  
 A:Reference number: A91586; MUID:86167830; PMID:3443302  
 A:Accession: A27363  
 A:Molecule type: mRNA  
 A:Residues: 1-76 <KN2>  
 A:Cross-references: GB:M19160  
 R:Henthorn, P.S.; Knoll, B.J.; Raducha, M.; Rothblum, K.N.; Slaughter, C.; Weiss, M.; La  
 Proc. Natl. Acad. Sci. U.S.A. 83, 5597-5601, 1986  
 A:Title: Products of two common alleles at the locus for human placental alkaline phosph  
 A:Reference number: A94113; MUID:86287303; PMID:3461452  
 A:Accession: A24318  
 A:Molecule type: mRNA  
 A:Residues: 23-535 <HEN>  
 A:Cross-references: GB:M14169  
 A:Note: parts of this sequence, including the amino end of the mature protein, were dete  
 A:Note: this is polymorphism type 1; 67-Leu also found  
 A:Accession: B24318  
 A:Molecule type: mRNA  
 A:Residues: 1-24, 'L', 26-65, 'V', 67-262, 'H', 264-276, 'R', 278-284, 'A', 286-388, 'C', 390-393, 'G  
 A:Cross-references: GB:M4170; NID:g178469; PIDN:AAA51709.1; PID:g178470  
 A:Note: parts of this sequence, including the amino end of the mature protein, were dete  
 A:Note: this is polymorphism type 3; 89-Leu also found  
 R:Millan, J.L.  
 J. Biol. Chem. 261, 3112-3115, 1986  
 A:Title: Molecular cloning and sequence analysis of human placental alkaline phosphatase  
 A:Reference number: A25385; MUID:86140079; PMID:3512548  
 A:Accession: A25385  
 A:Molecule type: mRNA  
 A:Residues: 1-230, 'P', 232-535 <MIL>  
 A:Cross-references: GB:M13077; NID:g178471; PIDN:ACC97139.1; PID:g178474

R:Ovitt, C.E.; Strauss, A.W.; Alpers, D.H.; Chou, J.Y.; Boine, I.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 3781-3785, 1986  
 A:Title: Expression of different-sized placental alkaline phosphatase mRNAs in plac  
 A:Reference number: A94099; MUID:86233318; PMID:3459156  
 A:Accession: A94099  
 A:Molecule type: mRNA  
 A:Residues: 382-435, 'T', 437-535 <OVI>  
 A:Note: a soluble or serum form was also found that may arise by proteolytic cleavag  
 R:Kam, W.; Clauser, E.; Kim, Y.S.; Kan, Y.W.; Rutter, W.J.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 8715-8719, 1985  
 A:Title: Cloning, sequencing, and chromosomal localization of human term placental al  
 A:Reference number: A94082; MUID:86094295; PMID:3001717  
 A:Accession: A94082  
 A:Molecule type: mRNA  
 A:Residues: 6-24, 'L', 26-260, 'GE', 263-323, 'H', 325-395, 'FI', 398-400, 'A', 402-535 <KAM>  
 A:Cross-references: GB:M1251; NID:g178463; PIDN:AAA51706.1; PID:g178464  
 R:McAnovic, R.; Bailey, C.A.; Brink, L.; Gerber, L.; Pan, Y.C.E.; Hulmes, J.D.; Ude  
 Proc. Natl. Acad. Sci. U.S.A. 85, 1398-1402, 1988  
 A:Title: Aspartic acid-484 of nascent placental alkaline phosphatase condenses with  
 A:Reference number: A28958; MUID:88144444; PMID:3422741  
 A:Accession: A28958  
 A:Molecule type: Protein  
 A:Residues: 485-535 <MIC>  
 R:Ezra, E.; Blacher, R.; Udenfriend, S.  
 Biochem. Biophys. Res. Commun. 116, 1076-1083, 1983  
 A:Title: Purification and partial sequencing of human placental alkaline phosphatase  
 A:Reference number: A90114; MUID:84079906; PMID:6651840  
 A:Accession: A90114  
 A:Molecule type: Protein  
 A:Residues: 23-64 <EKR>  
 C:Comment: This is the form of the enzyme predominantly expressed in the placenta aft  
 C:Genetics:  
 A:Gene: GDB:ALPP; ALP1  
 A:Cross-references: GDB:119672; OMIM:171800  
 A:Map position: 2q37.1-2q37.1  
 A:Insertions: 26/1; 65/1; 103/3; 162/1; 219/3; 264/3; 289/1; 334/1; 398/1; 437/1  
 A:Note: this gene is very polymorphic  
 C:Complex: homodimer  
 C:Function:  
 A:Description: catalyzes the hydrolysis of phosphate monoesters  
 C:Superfamily: alkaline phosphatase  
 C:Keywords: blocked carboxyl end; glycoprotein; homodimer; lipoprotein; membrane prot  
 F:1-22/Domain: signal sequence #status predicted <SIG>  
 F:23-506/Product: alkaline phosphatase, placental #status experimental <MAT>  
 F:507-535/Domain: carboxyl-terminal propeptide #status predicted <CTP>  
 F:114/Active site: Ser (phosphoserine intermediate) #status predicted  
 F:144/271/Binding site: carboxylate (Asn) (covalent) #status predicted  
 F:188/Binding site: substrate phosphate (Arg) #status predicted  
 F:342,382,454/Binding site: zinc (His) #status predicted  
 F:506/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Asp) (in mature t

Query Match 7.1%; Score 82; DB 1; Length 535;  
 Best Local Similarity 27.9%; Pred. No. 24;  
 Matches 43; Conservative 16; Mismatches 45; Indels 50; Gaps 8;

OY 76 KAPRRATAHYEV---HPRFG---QGAQAGVDGTVSGWEETKINSSEPLRDROIGEF 129  
 403 KAPRRKA---YTVLLYGNGGVYLDKGDARDYTESGSEFYRQSAVPLDETHAGED 458  
 OY 130 TVIRAGLYLYLCQVHPEBGKAVYIKLDLVNGV-----LALNCLEFSN----- 173  
 459 VAV-----FARGQAR-----LVHGVDEQTFIAHVMFAACLEBYTRACDLAPP 501  
 OY 174 ---TAASPGPOLRLCOVNSGLPLRPSSLRIRT 204  
 502 AGTTDAHPRGSV---VPALPLLAAGTLULLER 531

Search completed: May 8, 2003, 02:22:03  
 Job time : 15.3772 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## SUMMARIES

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 8, 2003, 02:22:47 ; Search time 166.65 Seconds

(without alignments)  
3040.496 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162

Sequence: 1 VLSIGLALACTGLLVVSL.....PMAHLKAPFLTYGFLQVH 225

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xip  
-Q/cgnt2\_1/USFTO.spool/US09245198/runat\_06052003\_170239\_9605/app.query.fasta\_1.846  
-DB=N-Geneseq\_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pcr -THR.MAX=100 -THR.MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRR=US09245198.0CGN.1.1.323 @runat.06052003\_170239\_9605 -NCPU=6 -ICPU=3  
-NO.XLPPY -NO.MMAP -LARGEPOURTY -NEG.SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N-Geneseq\_101002:\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Match	Length	DB	ID	Description
1	1162	100.0	1168	19	AAV18599	Mus musculus tumour
2	1089	93.7	701	20	AAV23425	Mouse TRRL3 DNA.
3	1020	87.8	1236	19	AAV47613	TNF related endoth
4	1020	87.8	1236	22	AAD04350	Human TREPA (TNF r
5	1020	87.8	1353	21	AAV49717	Human PRO207 CDNA
6	1020	87.8	1353	24	ABK40255	CDNA encoding huma
7	1020	87.8	1364	24	ABK43481	Human CDNA encodin
8	1020	87.8	1373	19	AAV18600	Homo sapiens tumou
9	1020	87.8	1421	20	AAV56000	Human tumour necro
10	951	81.8	898	22	AAV03964	Expression vector
11	951	81.8	1030	20	AAV23424	Human TRRL3 DNA.
12	113.5	9.8	2148	24	ABK11680	DNA encoding tumou
13	111	9.6	1221	24	ABK11679	DNA encoding tumou
14	110.5	9.5	978	23	ABK11473	Drosophila melanog
15	106.5	9.2	1630	19	AAV41377	NF-kB receptor act
16	106.5	9.2	1630	19	AAV41371	NF-kB receptor act
17	106.5	9.2	1630	22	AAD15310	Murine receptor ac
18	106.5	9.2	1630	22	AAD08714	Murine RANKL (reco
19	106.5	9.2	1630	22	AAD05903	Human secreted pro
20	106.5	9.2	3878	22	AAD05230	Human prosate CDNA
21	106.5	9.2	4034	22	AAV54040	Human prosate CDNA
22	106.5	9.2	4034	22	AAV53868	P533S CDNA splice
23	106.5	9.2	4034	24	ABK195411	Human P533S splice
24	106.5	9.2	4894	22	AAV54038	Human prosate CDNA
25	106.5	9.2	4894	22	AAV53866	P533S CDNA splice
26	106.5	9.2	4894	24	ABK195409	Human P533S splice
27	106.5	9.2	6976	22	AAV54041	Human prosate CDNA
28	106.5	9.2	6976	22	AAV53869	P533S CDNA splice
29	106.5	9.2	6976	24	ABK195412	Human P533S splice
30	105.5	9.1	81800	22	ABK84756	Human CDNA differe
31	103	8.9	684	22	AAV62935	HIV-1 gp120 V3 loo
32	102.5	8.8	951	19	AAV69960	Nucleotide sequenc
33	102.5	8.8	951	21	AAV39965	Mouse OSM nucleoti
34	102.5	8.8	951	21	AAV39965	DNA encoding a mur
35	102.5	8.8	951	21	AAV49024	Nucleic acid format
36	102.5	8.8	1538	19	AAV69886	Osteoclast format
37	102.5	8.8	2166	24	ABK11678	DNA encoding tumou
38	102.5	8.8	2191	19	AAV41489	Nucleotide sequenc
39	102.5	8.8	2237	20	AAV80224	Murine TRANCE enco
40	102.5	8.8	2237	24	ABK12880	CDNA encoding mous
41	102.5	8.8	2235	19	AAV70284	Human osteoprotege
42	102.5	8.8	2299	21	AAV29966	DNA encoding a mur
43	101.5	8.7	7332	24	ABK40768	Constitutively act
44	99.5	8.6	38186	20	AAV23028	Human MTH1 relate
45	99.5	8.6	38186	22	AAV30085	AC004449 CDNA clon

## ALIGNMENTS

RESULT 1  
AAV18599 standard; CDNA; 1168 BP.

AAV18599;

21-JUL-1998 (first entry)

Mus musculus tumour necrosis factor related ligand (TRELL) gene.

TRELL: tumour necrosis factor related ligand; tnfr; treatment;

KW cancer; autoimmune disease; immune system; stimulation; suppression;

KW graft rejection; ds.

OS Mus musculus.

XX Key Location/Qualifiers

FT CDS 2..679

```

FT      /tag- a
FT      /note- "tumour necrosis factor related ligand"
XX
XX      WO9805783-A1.
XX
XX      12-FEB-1998.
XX
XX      07-AUG-1997; 97WO-US13945.
XX
XX      18-MAR-1997; 97US-0040820.
XX      07-AUG-1996; 96US-0023541.
XX      18-OCT-1996; 96US-0028515.
XX
XX      (BIOC ) BIOGEN INC.
XX      (UYGE-) UNIV GENEVA FACULTY MEDICINE.
XX
XX      Browning JL, Chicheportliche Y;
XX
XX      WPI, 1998-145619/13.
XX      P-PSDB; AAW47524.
XX
XX      Tumour necrosis factor related ligand - useful for, e.g. treating
XX      cancer, auto-immune disease and immune responses to tissue grafts
XX
XX      Claim 2; Pages 45-46; 69pp; English.
XX
XX      The sequence is that encoding mouse tumour necrosis factor related
XX      ligand (TRELL). TRELL or active fragments can be included with a
XX      carrier in pharmaceutical compositions to treat cancer, autoimmune
XX      diseases or immune responses to tissue grafts, or to stimulate or
XX      suppress the immune system. It is useful to screen for TRELL
XX      receptors, by labelling with a detectable label and screening
XX      compositions for binding. Agents interfering with TRELL-receptor
XX      binding can also be screened for, can then be administered,
XX      optionally with interferon- gamma, to induce cell death or
XX      treat, suppress or alter immune responses (especially involving human
XX      adenocarcinoma cells) involving a signal pathway between TRELL and its
XX      receptor. The DNA sequence can be used in gene therapy for
XX      TRELL-related disorders in mammals (especially humans), e.g. tumours,
XX      autoimmune and inflammatory diseases or inherited genetic disorders,
XX      by introducing into cells, and expressing, therapeutically effective
XX      amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
XX      It may also be of use in the preparation of gene probes for
XX      screening natural/synthetic DNAs for TRELL-encoding sequences
XX      and for antisense therapy.
XX
XX      SQ Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.: 8 02e-97 Length: 1168
XX      Score: 1162.00 Matches: 225
XX      Percent Similarity: 100.00% Conservatve: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      DB: 19 Gaps: 0
XX
XX      US-09-245-198A-2 (1-225) x AAV18599 (1-1168)
XX
XX      1 ValIeuSerIeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValIeuSerIeu 20
XX      2 GTGCTAGAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 61
XX      21 GlySerTrpAlaThrIeuSerAlaGlnGluProSerGlnGluLeuThrAlaGluAsp 40
XX      62 GGGAGCTGGGGCAACGCTGCTGCCAGGAGCCTTCTCAGAGAGAGCTGCACAGAGAGAC 121
XX      41 ArgArgGluProGluLeuAsnProGlnThrGluGluSerGlnAspValAlaProPhe 60
XX      122 CGCGGGAGACCCCTCACTGAATCCCAAGAGAGAAAGCAAGATGGTACCTTTC 181
XX      61 LeuGluGlnLeuValArgProArgSerAlaProIleGlyArgIleValAlaArgProArg 80
XX      182 TTGGAAACACTAGTCCGGGCTCGAAGAAAGTCTCTAAAGCCGGAGAGCGGCTCGC 241

```

```

XX      81 ArgAlaIleAlaAlaHisIleTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAla 100
XX      242 CGAGCTATTGACGCCCTTTATGAGCTTCATCTCGGCAGAGACAGATGAGCAAGCA 301
XX      101 GlyValAspGlyThrValSerGlyTyrPgluGluThrIleAsnSerSerSerProLeu 120
XX      302 GGTGTGATGGAGACAGTGAAGTGGTGGGAAGACCAAAATCAACAGCTCCACCCCTCG 361
XX      121 ArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrIleTyr 140
XX      362 CGGTACGACCCGACAGATGGGGAATTTACAGTATCAGGGCTGGGCTTACTACTGTAC 421
XX      141 CysGlnValHisPheAspGlnGlyIleValValIleTyrLeuIleAsnSerSerProLeu 160
XX      422 TGTGAGGTGACCTTATGAGGAAAGCTGTCTACTGAAAGCTGACTGCTGGTGAAC 481
XX      161 GlyValIleuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerProGly 180
XX      482 GGTGTGCTGGCCCTGGGCTGCTGGAAGATTCACAGCCACAGACAGCAAGCTCTCGGG 541
XX      181 ProGlnLeuArgIleuGlyValSerGlyLeuLeuProIleuArgProGlySerSerIeu 200
XX      542 CCCAGCTCCGTTGTGCTGAGGTGCTGGGCTGTGCCGCTGGCGGAGGCTTCCCTT 601
XX      201 ArgIleArgThrIeuProTrpAlaHisIleuValAlaAlaProPheLeuThrTyrPheGly 220
XX      602 CGGATCGGACCCCTCCCTGGGCTGATCTTAAGGCTGCCCCCTTCCTAACCTTTCGGA 661
XX      221 LeuPheGlnValHis 225
XX      662 CTCTTCAAGTTCAC 676
XX
XX      RESULT 2
XX      AAX23425
XX      ID AAX23425 standard; DNA; 701 BP.
XX      AC AAX23425;
XX      DT 18-JUN-1999 (first entry)
XX      DE Mouse TNRL3 DNA.
XX      KW Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
XX      developmental abnormality; gestational abnormality; prostate cancer;
XX      APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
XX      cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
XX      apoptosis; mouse; ss.
XX      OS Mus sp.
XX      FH Key Location/Qualifiers
XX      FT CDS 1..636
XX      FT /tag- a
XX      FT /product- "TNRL3"
XX
XX      WO911791-A2.
XX      11-MAR-1999.
XX      04-SEP-1998; 98WO-US18393.
XX      05-SEP-1997; 97US-0924634.
XX      (UNITW ) UNIV WASHINGTON.
XX      Chaudhary PM;
XX      WPI; 1999-205191/17.
XX      P-PSDB; AAW93591.
XX      New Tumour Necrosis Factor family receptor polypeptides and ligands -
XX      useful for diagnosis and treatment of prostate cancer and

```

developmental or gestational abnormalities

XX Example VII; Fig 13B; 156pp; English.

XX This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: ApO4, ApO6, ApO8 and ApO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. ApO4 is useful for diagnosing prostate cancer  
 CC by determining levels of ApO4 in an individual. Prostate cancer can also  
 CC be treated using ApO4 selective binding agents linked to a therapeutic  
 CC moiety. ApO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. ApO4 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the change in ApO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using ApO4 polypeptides/active  
 CC fragments and ApO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of ApO4 and detecting a change in level of ApO4  
 CC activity. The method is performed in vivo or in vitro. ApO polypeptides  
 CC are all useful as immunogens for preparing antibodies. ApO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. ApO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.

XX Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;

Alignment Scores:

Pred. No.: 1,94e-90 Length: 701  
 Score: 1089.00 Matches: 210  
 Percent Similarity: 99.53% Conservative: 0  
 Best Local Similarity: 99.53% Mismatches: 1  
 Query Match: 93.72% Indels: 0  
 DB: Gaps: 0

US-09-245-198a-2 (1-225) x AAX23425 (1-701)

QY 15 LeuValValValSerIeuGlySerTrpAlaThrIeuSerAlaGlnIuProSerGlnIu 34  
 DB 1 CTGGTCTGCTGACCTGGGAGAGCTGGCAACGCTGTCTGCCAGGACCTTCTCAGAG 60  
 QY 35 GluLeuThrAlaGluAspArgArgGluProProGluIeuAsnProGlnThrGlnIuSer 54  
 DB 61 GAGCTGACAGCAGAGAGAGCCGGGAGGCCCTGTAACGATCCCAACAGAGAGAAAGC 120  
 QY 55 GlnAspValValProIleuGlnIuGlnIeuValArgProArgSerAlaProIuGly 74  
 DB 121 CAGGATGTGTACTTCTTCTGGAACAACCTAGTCCGCCCTCGAAGAGTCTCTAAAGGC 180  
 QY 75 ArgIysAlaArgProArgArgAlaIleAlaIleHisTrpGluValHisProArgProIy 94  
 DB 181 CGGAGGCGCGGCTCGGCGAGCTATGTCACGCCATTTATGAGTTCACTCGGCCAGGA 240  
 QY 95 GlnAspGlnAlaGlnAlaGlnValAspIuIuValSerGlyTrpGlnIuThrLysIle 114  
 DB 241 CAGATGAGAGCAAGAGAGGTGTGATGGAGACAGTGGCTGGGAAGACCAAAATC 300  
 QY 115 AsnSerSerProIeuArgTrpAspArgGlnIleGlyIuPheThrValIleArgAla 134  
 DB 301 AACGCTCCACCCCTCTGCTGACGACCGCAGATGGGGAATTTACAGTCATCAGGGCT 360  
 QY 135 GlyLeuTrpTrpLeuTrpCysGlnValHisPheAspGlnIuGlyLysAlaValTrpLeuLys 154  
 DB 361 GGGCTTACTACTCTGACTGTCAGGTGACATTTGTATGAGGAAAGGCTGTACTCTGAG 420  
 QY 155 LeuAspLeuValAsnGlyValIleuAlaLeuArgCysIeuGlnIuPheSerAlaThr 174  
 DB 421 CTGAGCTTGTGCTGAGAGGTGTGCTGGCCCTGGCTGCCGGAAGATTTTCAGCCCA 480  
 QY 175 AlaAlaSerSerProGlyProGlnIeuArgLysGlnValSerGlyLeuLeuProIeu 194  
 DB 481 GCAGCAAGCTCTCTCTGGGCGCCAGCTCGTTGTGTGCCAGGTGTGTGGCTGTGGCGCTG 540

QY 195 ArpProGlySerSerIeuArgIleArpThrIeuProTrpAlaHisIleuLysAlaAlaPro 214  
 DB 541 CGGCGAGGGGTCTTCCCTTGGAGTCGACCCCTCCCTGGGCTCATTTAAAGCTGCCCC 600

QY 215 PheLeuThrTrpPheGlyLeuPheGlnValHis 225  
 DB 601 TTCCTAACCTACTTTGGACCTTTCAGACTTTCAGACTTTCAG 633

RESULT 3

AAV47613 standard; cDNA; 1236 BP.

AAV47613;

27-OCT-1998 (first entry)

TNF related endothelium proliferative agent gene.

ss; TNF: endothelium proliferative agent; TREPA; wound healing; cancer; tissue grafting; vascularisation; apoptosis; autolimmune; birth control.

Homo sapiens.

Key Location/Qualifiers

FT CDS 1..750

FT /tag= a

FT /product= "TREPA"

MO9835061-A2.

13-AUG-1998.

12-FEB-1998; 98MO-US02859.

10-FEB-1998; 98US-0021706.

12-FEB-1997; 97US-0798692.

(ABBO ) ABBOTT LAB.

Wiley SR;

WPI: 1998-447255/38.

P-PSDB; AAM29745.

Detecting nucleic acid encoding TREPA - useful for diagnosis and treatment of autoimmune disease, tumours and inflammation

Claim 11; Page 123-4; 142pp; English.

CC The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TRPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).

Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

Alignment Scores:

Pred. No.: 7.98e-84 Length: 1236  
 Score: 1020.00 Matches: 199  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 88.84% Mismatches: 16

```

Query Match:      87.78%      Indels:      0
DB:               19          Gaps:         0
US-09-245-198a-2 (1-225) x AAV47613 (1-1236)

OY      2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      76 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 135

OY      22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGlnAspArg 41
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      136 AGCGGGGACATGCTGCTGGCCAGAGACCTGGCCAGAGAGAGAGAGAGAGAGAGAC 195

OY      42 ArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPheLeu 61
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      196 CAGAGACCCGCGGACATGATCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 255

OY      62 GluGlnLeuValArgProArgArgSerAlaProGlyGlyArgGlyAlaArgProArgArg 81
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      256 AACCGACTAGTTCGGCTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 315

OY      82 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      316 GCGATGCGAGCCCATTAATGAGTTCAATCCAGAGCTGGAGAGAGAGAGAGAGAGAGAG 375

OY      102 ValAspGlyThrValSerGlyTyrGluGluThrLysIleAsnSerSerSerProLeuArg 121
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      376 GTGAGACGGGACATGATGCTGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 435

OY      122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      436 TACACCGGCGAGATCGGGGAGTTATATACACCGGCTGGCTGCTACTACTACTACT 495

OY      142 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      496 CAGGTGACATTTATAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 555

OY      162 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAsnSerSerProLeuArg 181
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      556 GTGCTGGCCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 615

OY      182 GlnLeuArgLeuGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      616 CAGCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 675

OY      202 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      676 ATCGGACCCCTCCCGCGGCGCCATCTCAAGAGCTGCCCTTCTCACTTCCGAGCTC 735

OY      222 PheGlnValHis 225
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      736 TTCAGAGTTCAC 747

RESULT 4
ID      AAD04350 standard; cDNA: 1236 BP.
AC      AAD04350;
XX
XX
XX      04-JUL-2001 (first entry)
DE      Human TREPA (TNF related endothelium proliferative agent) cDNA.
XX
XX      Human: tumour necrosis factor; TNF; angiogenesis; wound healing;
KM      TREPA; TNF related endothelium proliferative agent; tumour; metastasis;
KM      grafting; vulnerability; ss.
XX
XX      Homo sapiens.
XX
XX      Key      Location/Qualifiers
FH      CDS      1..750
FT
FT      /product= "Human TREPA (TNF related endothelium

```

```

FT      FT      proliferative agent)"
XX      XX
PN      US6207642-B1.
XX      XX
XX      27-MAR-2001.
XX      XX
XX      26-JUN-1998; 98US-0105343.
XX      XX
XX      12-FEB-1997; 97US-0798692.
XX      XX
XX      10-FEB-1998; 98US-0021706.
XX      XX
XX      (ABBO ) ABBOTT LAB.
XX      XX
XX      WILEY SR.
XX      XX
XX      WPI: 2001-280760/29.
XX      XX
XX      P-PSDB; AAE00891.
XX      XX
XX      Inducing angiogenesis in mammal at desired sites for promoting wound
XX      healing, by administering soluble fragment of extracellular domain of
XX      tumor necrosis factor related endothelium proliferative agent protein
XX      -
XX      Example 2; Column 73-74; 53pp; English.
XX      XX
XX      The present invention relates to extracellular signal molecules,
XX      particularly members of tumour necrosis factor (TNF) family molecules
XX      designated as TREPA (TNF related endothelium proliferative agent).
XX      Soluble biologically active TREPA are used to treat TREPA-associated
XX      diseases, tumours or metastases. TREPA is used for inducing angiogenesis
XX      in human for promoting wound healing and for vascularising grafted
XX      tissue for successful grafting and to promote tissue grafts.
XX      The present sequence is a cDNA clone ID #690050 encoding human TREPA.
XX      XX
SQ      Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other:

Alignment Scores:
Pred. No.:      7.98e-84      Length:      1236
Score:          1020.00      Matches:      199
Percent Similarity: 92.86%      Conservative: 9
Best Local Similarity: 88.84%      Mismatches: 16
Query Match:    87.78%      Indels:      0
DB:             22          Gaps:         0

US-09-245-198a-2 (1-225) x AAD04350 (1-1236)

OY      2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      76 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 135

OY      22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGlnAspArg 41
      ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      136 AGCGGGGACATGCTGCTGGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 195

OY      42 ArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPheLeu 61
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      196 CAGAGACCCGCGGACATGATCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 255

OY      62 GluGlnLeuValArgProArgArgSerAlaProGlyGlyArgGlyAlaArgProArgArg 81
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      256 AACCGACTAGTTCGGCTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 315

OY      82 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      316 GCGATGCGAGCCCATTAATGAGTTCAATCCAGAGCTGGAGAGAGAGAGAGAGAGAGAG 375

OY      102 ValAspGlyThrValSerGlyTyrGluGluThrLysIleAsnSerSerSerProLeuArg 121
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      376 GTGAGACGGGACATGATGCTGGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 435

OY      122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB      436 TACACCGGCGAGATCGGGGAGTTATATACACCGGCTGGCTGCTACTACTACTACT 495

```



QY 142 GlnValHisPheAspGluLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161  
 DB 496 CAGGTGCACCTTGGATGAGGGGAAAGCTGCTCTACCTGAAGCTGACTGCTGGTGGATGGT 555  
 QY 162 ValLeuAlaLeuArgCysLeuGluLuphSerAlaThrAlaLaserSerProGlyPro 181  
 DB 556 GTGCTGGCCCTGCTGCTGGATGAGAAATCTCAGCCACTGGCGGAGTTCCTCGGGGCC 615  
 QY 182 GlnLeuArgLeuGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201  
 DB 616 CAGCTCCGCGCTGCGAGGTGTCTGGGCTGTGGCCCTGCGCCAGGGCTCTCCCTGCGG 675  
 QY 202 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221  
 DB 676 ATCCGACACCTTCCTCGGGCCATCTCAAGGCTGCCCCCTTCTCACTTCCGAGCTC 735  
 QY 222 PheGlnValHis 225  
 DB 736 TTCAGGTTCAC 747  
 RESULT 5  
 AAA49717  
 ID AAA49717 standard; cDNA; 1353 BP.  
 AC AAA49717;  
 XX 25-SEP-2000 (first entry)  
 DE Human PRO207 cDNA clone DNA30879-1152.  
 KW PRO207; human; antitumour; tumour; therapy; cytostatic;  
 KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
 KW uterine cancer; prostate cancer; lung cancer; bladder cancer;  
 KW central nervous system cancer; melanoma; leukaemia; neoplasm; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 58..807  
 FT sig\_peptide 58..177  
 FT /\*tag= a  
 FT /\*tag= b  
 FT mat\_peptide 178..804  
 FT /\*tag= c  
 PN WO200037638-A2.  
 XX  
 PD 29-JUN-2000.  
 PF 02-DEC-1999; 99WO-US28565.  
 XX  
 PR 22-DEC-1998; 98US-0113296.  
 PR 08-MAR-1999; 99WO-US05028.  
 PR 21-APR-1999; 99US-0130232.  
 PR 28-APR-1999; 99US-0131445.  
 PR 14-MAY-1999; 99US-0134287.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 15-SEP-1999; 99WO-US21090.  
 PR 15-SEP-1999; 99WO-US21547.  
 XX  
 PA (GENE) GENENTECH INC.  
 PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;  
 PI Napier MA, Pittl RM, Wood WI;  
 XX  
 DR MPI: 2000-442668/38.  
 DR P-PSDB: AAY95338.  
 XX  
 PT Novel composition to inhibit neoplastic cell growth or for treating  
 PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,  
 PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or

PT PRO866 -  
 XX  
 PS Claim 20; Fig 3; 172pp; English.  
 XX  
 CC The present sequence is that of cDNA clone DNA30879-1152  
 CC (ATCC 209358) encoding human PRO207 (see AAY95338), which shows  
 CC homology to several members of the tumour necrosis factor family,  
 CC especially human lymphotoxin (23.4%). The cDNA was identified in a  
 CC foetal kidney cDNA library following identification of an expressed  
 CC sequence tag with homology to human Apo-2 ligand. A claimed method  
 CC for inhibiting the growth of a tumour cell comprises exposing the  
 CC tumour cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224,  
 CC PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866 (see  
 CC AAY95337-49), their agonists or chimeric polypeptides incorporating  
 CC them. The tumour is especially a cancer selected from breast,  
 CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and  
 CC central nervous system cancer, melanoma and leukaemia. Nucleic  
 CC acids encoding PRO179 etc. are used in the recombinant production  
 CC of the antitumour polypeptides.  
 XX  
 SQ Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other:  
 Alignment Scores:  
 Pred. No.: 8,95e-84 Length: 1353  
 Score: 1020.00 Matches: 199  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 88.84% Mismatches: 16  
 Query Match: 87.78% Indels: 0  
 DB: 21 Gaps: 0  
 US-09-245-198A-2 (1-225) x AAA49717 (1-1353)  
 QY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21  
 DB 133 CTGGGCGTGGGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGG 192  
 QY 22 SerTrpAlaThrLeuSerAlaGlnGluLuphSerGlnGluLeuThrAlaGlnAspArg 41  
 DB 193 ACCCGGGCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 252  
 QY 42 ArgGluProProGluLeuAsnProGlnThrGlnGluSerGlnAspValAProPheLeu 61  
 DB 253 CAGGACCGCTGCGAACTGAATCCCAAGACAGAAAGCAGAGATTCCTGCTTCCG 312  
 QY 62 GlnGlnLeuValAlaArgProArgArgSerAlaProPolsGlyArgGlyAlaArgProArg 81  
 DB 313 AACGACATGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 372  
 QY 82 AlaIleAlaAlaHisTyrGlnValHisProArgProGlnGlnAspGlyValAlaGlnAlaGly 101  
 DB 373 GCGATCGCAGCCCATTAATCAATTCATCCACACTGCGACAGAGCGGAGCGCAGGCT 432  
 QY 102 ValAspGlyThrValSerGlyTyrPgluGlnThrLysIleAsnSerSerSerProLeuArg 121  
 DB 433 GTGAGCGGACAGTAGTGTGGTGGGAGGAGCAATCAACAGCTCCAGCCCTGCGC 492  
 QY 122 TyrAspArgGlnIleGlyGlnPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141  
 DB 493 TGCACACCGCAGATCGGGAGTTATAGTCACCGGCTGCGCTCTACACTGACTGT 552  
 QY 142 GlnValHisPheAspGluLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161  
 DB 553 CAGGTGCACCTTGGATGAGGGGAAAGCTGCTCTACCTGAAGCTGACTGCTGGTGGATGGT 612  
 QY 162 ValLeuAlaLeuArgCysLeuGluLuphSerAlaThrAlaLaserSerProGlyPro 181  
 DB 613 GTGCTGGCCCTGCTGCTGGATGAGAAATCTCAGCCACTGGCGGAGTTCCTCGGGGCC 672  
 QY 182 GlnLeuArgLeuGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201  
 DB 673 CAGCTCCGCGCTGCGAGGTGTCTGGGCTGTGGCCCTGCGCCAGGGCTCTCCCTGCGG 732  
 QY 202 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221

Db	733	ATCCGACACCTCCCTGGGCCATCTCAGGCGTCCCTCTCCTACCTTGGA	792
Oy	222	PhcGlnVAlHts 225	
Db	793	TTCCAGGTTTCAc 804	
RESULT 6			
ABK40255			
ID	ABK40255	standard; cdna; 1353 BP.	
XX			
AC	ABK40255;		
XX			
DT	15-JUL-2002	(first entry)	
XX			
DE	cdna encoding human PRO207 polypeptide.		
XX			
KW	Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;		
KW	leukemia; neuronal disorder; stromal disorder; blastocoele disorder;		
KW	inflammatory disorder; immune disorder; angiogenic disorder;		
KW	gene therapy; cytostatic; neuroprotective; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200153486-A1.		
PD			
XX	26-JUL-2001.		
XX			
PF	11-FEB-2000; 2000WO-US03565.		
XX			
XX			
PR	08-MAR-1999; 99WO-US05028.		
PR	11-MAR-1999; 99US-123972P.		
PR	11-MAY-1999; 99US-133459P.		
PR	02-JUN-1999; 99WO-US12252.		
PR	22-JUN-1999; 99US-140650P.		
PR	22-JUN-1999; 99US-140653P.		
PR	20-JUL-1999; 99US-144758P.		
PR	26-JUL-1999; 99US-145698P.		
PR	28-JUL-1999; 99US-146272P.		
PR	17-AUG-1999; 99US-149385P.		
PR	31-AUG-1999; 99US-151689P.		
PR	01-SEP-1999; 99WO-US20111.		
PR	15-SEP-1999; 99WO-US21090.		
PR	30-NOV-1999; 99WO-US28313.		
PR	01-DEC-1999; 99WO-US28304.		
PR	01-DEC-1999; 99WO-US28634.		
PR	05-JAN-2000; 2000WO-US00219.		
XX			
PA	(GETH ) GENENTECH INC.		
XX			
PI	Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;		
PI	Marsers SA, Pan J, Pittl RM, Roy MA, Smith V, Stone DM;		
PI	Ketanabe CK, Wood WJ.		
XX			
DR	WPI; 2002-205567/26.		
XX	P-PSDB; AA086129.		
PT	Thirty five nucleic acids encoding PRO polypeptides, useful for		
PT	treating benign or malignant tumours, leukemias and lymphoid		
PT	malignancies, inflammatory, angiogenic and immunologic disorders -		
XX			
XX	Claim 50; Fig 3; 302pp; English.		
XX			
CC	The present invention relates to the isolation of novel human PRO		
CC	polypeptides and the polynucleotide sequences encoding them. The		
CC	PRO polypeptides, agonists, antagonists or anti-PRO antibodies are		
CC	useful for treating benign or malignant tumours (e.g. renal, kidney,		
CC	bladder, breast, etc), leukemias and lymphoid malignancies, other		
CC	disorders such as neuronal, glial, astrocytal, hypothalamic, glandular		
CC	macrophagal, stromal and blastocoele disorders, inflammatory, immune		
CC	and angiogenic disorders. The polynucleotide sequences are also		
CC	useful in gene therapy. ABK40254-ABK40288 encode for the human PRO		
CC	polypeptides of the invention.		

[illegible]



DR P-PSDB; AAN47525.

XX Tumour necrosis factor related ligand - useful for, e.g. treating  
PT cancer, auto-immune disease and immune responses to tissue grafts  
XX  
XX Claim 2; Pages 48-50; 69pp; English.

XX The sequence is that encoding human tumour necrosis factor related  
CC ligand (TRELL). TRELL or active fragments can be included with a  
CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
CC diseases or immune responses to tissue grafts, or to stimulate or  
CC suppress the immune system. It is useful to screen for TRELL  
CC receptors, by labelling with a detectable label and screening  
CC compositions for binding. Agents interfering with TRELL-receptor  
CC binding can also be screened for, can then be administered,  
CC optionally with interferon-gamma, to induce cell death or  
CC treat, suppress or alter immune responses (especially involving human  
CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
CC receptor. The DNA sequence can be used in gene therapy for  
CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
CC autoimmune and inflammatory diseases or inherited genetic disorders,  
CC by introducing into cells, and expressing, therapeutically effective  
CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
CC It may also be of use in the preparation of prepare probes for  
CC screening natural/synthetic DNAs for TRELL-encoding sequences  
CC and for antisense therapy.

SO Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other:

Alignment Scores:

Pred. No.:	9.12e-84	Length:	1373
Score:	1020.00	Matches:	199
Percent Similarity:	92.86%	Conservative:	9
Best Local Similarity:	88.84%	Mismatches:	16
Query Match:	87.78%	Indels:	0
DB:	19	Gaps:	0

US-09-245-198a-2 (1-225) x AAV18600 (1-1373)

```

OY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21
DB 101 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 240
OY 22 SerTPALaThrLeuSerAlaGlnGlnProSerGlnGlnGlnGlnGlnGlnGln 41
DB 241 AGCCGGGATCGCTCTCCGCCAGAGAGCTGCCAGAGAGCTGCCAGAGAGCTGCCAG 300
OY 42 ArgGluProProGluLeuAsnProGlnThrGlnGlnGlnGlnGlnGlnGln 61
DB 301 CAGGACCCGTCGAACTGAAATCCAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
OY 62 GlnGlnLeuValArgProArgArgSerAlaProGlyGlyGlyGlyGlyGlyGly 81
DB 361 AACCGACAGTGTGGCTCGCAGAGAGTCCACCTAAAGCCCGGAGAGAGAGAGAGAG 420
OY 82 AlaIleAlaAlaHisTyrGlyValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
DB 421 GCGATCGAGCCGATTAAGAGTTCATCCAGACTGAGACGAGAGAGAGAGAGAGAGAG 480
OY 102 ValAspGlyThrValSerGlyTTPGluGluThrValHisLeuSerSerProLeuArg 121
DB 481 GTGGACGGGACAGTGTGGCTCGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
OY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141
DB 541 TACAACCGCCGATCGGGAGATTATAGTACCCGGGCTGGGCTTACTACTACTACTACT 600
OY 142 GlnValHisPheAspGlnGlyValAlaValTyrLeuValLeuAspLeuLeuValAsnGly 161
DB 601 CAGGTGACCTTTGATGAGGGAGAGAGCTGTCTACTGAAAGCTGGAGCTGTGGAGAGGT 660
OY 162 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyPro 181
DB 181

```

```

DB 661 GTCTGCCCCCTGGCTGCTGGAGAGATTCACGCCATCGCGGAGTTCCTCGGGCCC 720
OY 182 GlnLeuArgLeuGlyGlnValSerGlyLeuLeuProLeuArgProGlySerLeuArg 201
DB 721 CAGCTCCGCTCTGGCCAGAGTGTCTGGCTGTGGCCCTGGCCGAGAGGTCTCCCTGGCG 780
OY 202 IleArgThrLeuProThrAlaHisLeuValAlaIleProPheLeuThrTyrPheGlyLeu 221
DB 781 ATCCGACCCCTCCCTGGGCGCCATCATCAAGCTGCCCTCTCTCCTCACTTCGAGACTC 840
OY 222 PheGlnValHis 225
DB 841 TTCACAGTTTAC 852

```

RESULT 9

AAK56000  
ID AAK56000 standard; DNA; 1421 BP.

AC AAK56000;

DT 15-JUL-1999 (first entry)

DE Human tumour necrosis factor Apo-3 ligand polynucleotide sequence.

DE Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;

KW NF-kappaB-dependent transcription; JNK/SAPK-dependent response;

KW cancer; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 92..841

FT /tag= a

FT /product= "Apo-3 ligand"

PN MO9919490-A1.

PD 22-APR-1999.

XX 09-OCT-1998; 98WO-US21407.

XX 17-DEC-1997; 97US-0069862.

XX 10-OCT-1997; 97US-0062037.

XX (GETH ) GENENTECH INC.

XX Ashkenazi AJ, Marsters SA, Pitti R;

XX WPI: 1999-287982/24.

XX P-PSDB; AAY09369.

XX New human Apo-3 ligand (a tumour necrosis factor) homologue

XX Claim 18; Fig 1; 74pp; English.

XX The present sequence encodes a human tumour necrosis factor (TNF) and

XX lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has

XX cyostatic activity. Apo-3 ligand can be used to induce apoptosis in

XX mammalian cancer cells; to induce NF-kappaB-dependent transcription and

XX to induce JNK/SAPK-dependent responses in mammalian cells.

SO Sequence 1421 BP; 281 A; 464 C; 404 G; 272 T; 0 other:

Alignment Scores:

Pred. No.:	9.53e-84	Length:	1421
Score:	1020.00	Matches:	199
Percent Similarity:	92.86%	Conservative:	9
Best Local Similarity:	88.84%	Mismatches:	16
Query Match:	87.78%	Indels:	0
DB:	20	Gaps:	0

US-09-245-198a-2 (1-225) x AAK56000 (1-1421)



```

QY 59 ProPheLeuGlInLeuValArgProArgArgSerAlaProLysGlyAlaArg 78
DB 370 CCTTCTGTAACCGACTAGTTCGGCTCCGACAGAGTCAACGCGGAAACACGG 429
QY 79 ProArgAlaAlaIleAlaHisTyrGluValHisProArgProGlyAla 98
DB 430 GCTCGAGAGCCGATCGCACCCCTTATGAAATTCATCCAGCCTGACGACGAGCG 489
QY 99 GlnAlaGlyValAlaSpGlyThrValSerGlyTyrGluGluThrValIleAsnSerSer 118
DB 490 CAGGAGGTGTGGACGGAGCAGTGTGCTGGAGGAAAGCCAGAAATCAACAGCTCCACG 549
QY 119 ProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyr 138
DB 550 CCTCTGCGCTACAAACCGCCAGATCGGGGAGTTATAGTCAACCGGGCTGGGCTCTAC 609
QY 139 LeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu 158
DB 610 CTGTACTGTCAAGTGCATCTTATGATGAGGGAAGGCTGTCTACCTGAACTGAGCTTG 669
QY 159 ValAsnGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleSer 178
DB 670 GTGATGTGTGTCTGGCCCTGGCGCTGGAGGAATTCACGACCTGCGGCACTTCC 729
QY 179 ProGlyProGlnLeuArgLysGlnValSerGlyLeuLeuProLeuArgProGlySer 198
DB 730 CTCGGGCCCCAGCTCCGCTCTGCGACAGGTCTGTGGGCTGTGGCCCTGGCGGCACTTCC 789
QY 199 SerLeuArgGlyLeuArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThr 218
DB 790 TCCCTGGGAGATCCGACACCTCCCTGGGCGCCATCTCAAGGCTGCCCTTCTCACTTAC 849
QY 219 PheGlyLeuPheGlnValHis 225
DB 850 TTCGAGACTCTCCAGGTTCCAC 870
RESULT 11
AAK23424
ID AAK23424 standard; DNA; 1030 BP.
XX
AC AAK23424;
XX
DT 18-JUN-1999 (first entry)
XX
DE Human TNRL3 DNA.
XX
KW Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW developmental abnormality; gestational abnormality; prostate cancer;
KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW apoptosis; human; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..627
FT /tag= a
FT /product= "TNRL3"
XX
PD MO9911791-A2.
XX
PD 11-MAR-1999.
XX
PF 04-SEP-1998; 98MO-US18393.
XX
PR 05-SEP-1997; 97US-0924634.
XX
PA (UNITV) UNIV WASHINGTON.
XX
PI Chaudhary PM;
XX
DR WPI; 1999-205191/17.
DR P-PSDB; AAW93590.

```

```

XX
PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PT developmental or gestational abnormalities
XX
PS Example VII: Fig 13A; 156pp; English.
XX
CC This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the change in APO4
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
XX
SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other.
XX
Alignment Scores:
Pred. No.: 1.26e-77 Length: 1030
Score: 951.00 Matches: 185
Percent Similarity: 92.79% Conservative: 8
Best Local Similarity: 88.94% Mismatches: 15
Query Match: 81.84% Indels: 0
DB: Gaps: 0
US-09-245-198a-2 (1-225) x AAK23424 (1-1030)
QY 18 ValSerLeuGlySerTyrPalaThrLeuSerAlaGlnGluProSerGlnGluLeuThr 37
DB 1 GTCAAGTTGGGAGACCGGGGATCGCTGCTCCGCCAGAGGCTGCCAGGAGCTGTG 60
QY 38 AlaGluAspArgArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspVal 57
DB 61 GCAGAGGAGGACGACGAGCCCTCGGAAGTGAATCCAGACAGAAAGCCAGAGATCT 120
QY 58 ValProPheLeuGluGlnLeuValArgProArgArgSerAlaProLysGlyAla 77
DB 121 GCGCTTCTCGAACCACTAGTTGCGCTCGCAGAAAGTGCACCTAAAGCGCGGAAAAA 180
QY 78 ArgProArgArgAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGly 97
DB 181 CGGGCTCGAAGAGGATCGACGCCCATTTAGTAATTCACACCTGGACGAGCGA 240
QY 98 AlaGlnAlaGlyValAlaSpGlyThrValSerGlyTyrGluGluThrValIleAsnSer 117
DB 241 GCGGAGGAGGATGTGGAGGAGCAGTGAAGTGTGGAGGAGGAGCAACATCAACAGCTCC 300
QY 118 SerProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyr 137
DB 301 AGCCCTCTGCGTACTACCCGCAAGTCCGGGAGTTATAGTCAACCGGGCTGGGCTTAC 360
QY 138 TyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu 157
DB 361 TACTGTACTGTCAAGGTGCATTTGATGAGGGAAGGCTGTCTTACCTGAAGCTGGACTTG 420
QY 158 LeuValAsnGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleSer 177
DB 421 CTGTGTGATGTGTCTGGGCGCTGCGCTGAGGAATTCACAGCACTGCGGCGACGT 480

```

```

OY 178 SerProGlyProGlnLeuArgLeuGlyCysGlnValSerGlyLeuLeuProLeuArgProGly 197
DB 481 TCCCTCGGGGGCCCGACCTCGCTTGCAGAGTGTGTGGCTTGGCCCTGGGGCCAGGG 540
OY 198 SerSerLeuArgGlyLeuArgProGlnLeuArgLeuGlyCysGlnValSerGlyLeuLeuProLeuArgProGly 217
DB 541 TCCCTCGGGGGCCCGACCTCGCTTGCAGAGTGTGTGGCTTGGCCCTGGGGCCAGGG 600
OY 218 TyrPheGlyLeuPheGlnValHis 225
DB 601 TACTTCGGAGCTTCCTCCAGGTTCCAC 624

RESULT 12
ABK11680
ID ABK11680 standard; DNA: 2148 BP.
AC ABK11680;
DE 05-JUN-2002 (first entry)
XX DNA encoding tumour necrosis factor variant 2 (TNFV2).
XX Tumour necrosis factor; TNF; pesticide; crop-damaging insect;
XX farm animal insect; epithelial morphogenesis; cell-matrix adhesion;
XX X-linked anhidrotic (hypohidrotic) ectodermal dysplasia;
XX X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;
XX sparse hair; sweat gland aberration; endotoxic shock; inflammation;
XX haemorrhagic necrosis of tumour; cytotoxicity; TNFV2;
XX obesity-linked insulin resistance; gene; ds.
XX Drosophila melanogaster.
XX
XX Key Location/Qualifiers
FH sig_peptide 634..789
FT /*tag= a
FT CDS 634..1863
FT /*tag= b
FT /product= "TNFV1"
FT /note= "Tumour necrosis factor variant 1"
FT mat_peptide 790..1860
FT /*tag= c
FT /label= mature_TNFV1
FT misc_difference 634..1860
FT /*tag= d
FT /note= "Specifically claimed in claim 22"
XX
XX US2002012968-A1.
XX
XX 31-JAN-2002.
XX
XX 20-MAR-2001; 2001US-0813329.
XX
XX 21-MAR-2000; 2000US-190816P.
XX
XX (CARR/) CARROLL P M.
XX (CHEN/) CHEN J.
XX (RAMA/) RAMANATHAN C S.
XX (XIAO/) XIAO H.
XX (GUAN/) GUAN B.
XX (BOWE/) BOWEN M A.
XX
XX Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA.
XX WPT: 2002-195121/25.
XX P-PSDB: AAU77718.
XX
XX New Drosophila tumour necrosis factor molecule, useful in controlling
XX agriculturally important pests, e.g. comprises modifying the growth,
XX feeding or reproduction of crop-damaging insects or insects of farm
XX animals -
XX Claim 2; Fig 3A-C; 119pp; English.
XX

```

```

CC The invention describes an isolated tumour necrosis factor polypeptide
CC (TNF). The polypeptide and polynucleotide are useful in controlling
CC agriculturally important pests, particularly by modifying the growth,
CC feeding and/or reproduction of crop-damaging insects or insects of farm
CC animals. The polypeptide and polynucleotide are useful for modulating
CC epithelial morphogenesis, cell-matrix adhesion in flies and mammals.
CC Thus the polypeptide and polynucleotide may be useful for treating,
CC ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal
CC dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal
CC dysplasia-like disorders, e.g. sparse hair, abnormal or missing teeth or
CC sweat gland aberrations in animals (e.g. insects and potentially humans),
CC endotoxic shock, inflammation, haemorrhagic necrosis of tumours,
CC cytotoxicity and obesity-linked insulin resistance, all of which involve
CC TNF molecules. This sequence encodes the drosophila melanogaster tumour
CC necrosis factor variant 2 (TNFV2) protein, described in the invention.
XX
XX SQ Sequence 2148 BP; 645 A; 510 C; 506 G; 487 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.962 Length: 2148
XX Score: 113.50 Matches: 50
XX Percent Similarity: 41.54% Conservative: 31
XX Best Local Similarity: 25.64% Mismatches: 78
XX Query Match: 9.77% Indels: 36
XX DB: Gaps: 9
XX
XX US-09-245-198a-2 (1-225) x ABK11680 (1-2148)
OY 41 ArgArgGluProProGlnLeuArgLeuGlyCysGlnValSerGlyLeuLeuProLeuArgProGly 60
DB 1353 CAAGAGAGAGCCCTGCACCTTCCACCA-----CCGTCCAGAAATCATCTCCGCCCA 1403
OY 60 eLeuGlnGlnLeuValArgProArgSerGlyLeuValArgProArg 80
DB 1404 TCGCCACCTCTAGTCCGAAAGCCAGATCC-----GAGGACTCCAGCCA-- 1449
OY 80 gArgAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGlyValGlnAl 100
DB 1450 -----GCAGCCCATTTCCACTTGCAGCAGCGCGCCCTCCACCAAGAAATG-- 1497
OY 100 aGlyValAspGlyThrValSerGlyTyrPgluThrLysIleAsnSerSerProle 120
DB 1498 -GGTACCATGAGATATGATATGATAGAAATGATAAAGCAAACTTTATCAGGGA-- 1554
OY 120 uArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeu 140
DB 1555 -CACTTCAACGCGCGATGGCTCTTGACGGTACCAATACAGCGCTATATTACGTATA 1613
OY 140 rCysGlnValHisPheAspGluGlyLysAlaValIleuLysLeuAspLeuValAs 160
DB 1614 CGCCCAAGATATGCTACACAACTCCGACGAC-----CAGCA 1649
OY 160 nGlyValLeuAla-----LeuArgCysLeuGluGluPheSerAl 173
DB 1650 CGGATTTATCGTCTTCAAGAGACACATCCATCTCGAGTGTCTGAAC----- 1698
OY 173 aThrAlaIleSerSerProGlyProGlnLeuArgLeuGlyCysGlnValSerGlyLeuLeu 193
DB 1699 -ACGGTCCCAACATGCGACATAGGTGACACCTCCACAGAGAGGTCTATCCA 1757
OY 193 oLeuArgProGlySerSerLeuArgIleArgThrLeu-----ProTyrAlaHisIste 210
DB 1758 CCTGGACGAAACGAGAGATCCATCTGAAGACATTCACCAACGATCGCAATGCAATGCTCT 1817
OY 210 uLysAlaIleProPheLeuThrTyrPheGlyLeuPheGlnVal 224
DB 1818 GCGGGAGGGAACCAACCGAAAGCTACTTGTCATCTCAAGGTG 1860

RESULT 13
ABK11679
ID ABK11679 standard; DNA: 1221 BP.
AC ABK11679;

```

XX 05-JUN-2002 (first entry)

DT DNA encoding tumour necrosis factor variant 1 (TNFV1).

DE

XX

XX Tumour necrosis factor: TNF; pesticide; crop-damaging insect;

KW farm animal insect; epithelial morphogenesis; cell-matrix adhesion;

KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia;

KW X-linked anhidrotic (hypohidrotic) ectodermal dysplasia-like disorder;

KW sparse hair; sweat gland aberration; endotoxic shock; inflammation;

KW haemorrhagic necrosis of tumour; cytotoxicity; TNFV1;

KW obesity-linked insulin resistance; gene; ds.

OS Drosophila melanogaster.

XX

XX Key Location/Qualifiers

FT 1..156

FT sig\_peptide

FT /\*tag- a

FT 1..1221

FT CDS

FT /\*tag- b

FT /product- "TNFV1"

FT /note- "Tumour necrosis factor variant 1"

FT 157..1218

FT mat\_peptide

FT /\*tag- c

FT /label- mature\_TNFV1

FT 1..1218

FT misc\_difference

FT /\*tag- d

FT /note- "Specifically claimed in claim 18"

XX

XX US2002012968-A1.

XX

XX 31-JAN-2002.

XX

XX 20-MAR-2001; 2001US-0813329.

XX

XX 21-MAR-2000; 2000US-190816P.

XX

XX (CARR/) CARROLL P M.

XX (CHEN/) CHEN J.

XX (RAMA/) RAMANATHAN C S.

XX (XIAO/) XIAO H.

XX (GUAN/) GUAN B.

XX (BOWE/) BOWEN M A.

XX

XX Carroll PM, Chen J, Ramanathan CS, Xiao H, Guan B, Bowen MA;

XX WPL: 2002-195121/25.

XX

XX New Drosophila tumour necrosis factor molecule, useful in controlling

PT agriculturally important pests, e.g. comprises modifying the growth,

PT feeding or reproduction of crop-damaging insects or insects of farm

PT animals

XX

XX Claim 2; Fig 2A-B; 119p; English.

XX

XX The invention describes an isolated tumour necrosis factor polypeptide

CC (TNF). The polypeptide and polynucleotide are useful in controlling

CC agriculturally important pests, particularly by modifying the growth,

CC feeding and/or reproduction of crop-damaging insects or insects of farm

CC animals. The polypeptide and polynucleotide are useful for modulating

CC epithelial morphogenesis, cell-matrix adhesion in flies and mammals.

CC Thus the polypeptide and polynucleotide may be useful for treating,

CC ameliorating or preventing X-linked anhidrotic (hypohidrotic) ectodermal

CC dysplasia and X-linked anhidrotic (hypohidrotic) ectodermal

CC dysplasia-like disorders, e.g. sparse hair, abnormal or missing teeth or

CC sweat gland aberrations in animals (e.g. insects and potentially humans),

CC endotoxic shock, inflammation, haemorrhagic necrosis of tumours,

CC cytotoxicity and obesity-linked insulin resistance, all of which involve

CC TNF molecules. This sequence encodes the drosophila melanogaster tumour

CC necrosis factor variant 1 (TNFV1) protein, described in the invention.

XX

XX Sequence 1221 BP; 349 A; 323 C; 312 G; 237 T; 0 other;

Alignment Scores:

Pred. No.:	0.792	Length:	1221
Score:	111.00	Matches:	44
Percent Similarity:	39.81%	Conservative:	38
Best Local Similarity:	21.36%	Mismatches:	84
Query Match:	9.55%	Indels:	40
DB:	24	Gaps:	8

US-09-245-198a-2 (1-225) x ABL21473 (1-1221)

QY 29 GINGUPTROSERGLINGLULENTHRALAULSPARGALGULPROPTROGLULEASN 48

DB 691 CAGGAAAGATCTCAATCAAGGACCACTTCCAAAGAGAAATGATCTCCGCCATCGCCAC 750

QY 49 PROGINTHRGULUSERGLINASPVALAIPROPHLEAUGLINGLULEVALAIPROARG 68

DB 751 CTCCTAGTCCGCAAGGATGATCTCTT----- 780

QY 69 ARGSERALAPROLYSGIARGLYSALAARGPROARGALALALALALALALALALALALAL 88

DB 781 -----TCAGCCAGATCCGAGACTCGAGGCCA-----GCAGCCCATTTCCAC 822

QY 89 VALHISPROARGPROGLINGLINSPOLYALALALALALALALALALALALALALALALAL 108

DB 823 TTGACGACAGCGCGCGCTCACCAAGGAAATGTC---GGCTACCATGGAGATATGATACATA 879

QY 109 TRPGULINTHRLYSILEASNSERSESERPROLEUARGTYRRAAPARGLNLGLYGLU 128

DB 880 GGAATGATACGAGAGAACTTTATCAGGGA---CACTTTCAAGACGGGATGGCTC 936

QY 129 PHEATHVALIARGALAGLYLEUTYTYRLEUTYRCYSGINVALHISPHASPGULY 148

DB 937 TTGACGGTGACCAATACAGGCGCTATATACCTATACCCCGAGATATGCTACACACTCG 996

QY 149 LYSALVALYTYRLEUYLSLEASPLEULEVALNSGLYVALLEUALA----- 164

DB 997 CACGAC-----CAGACGATTTATCTGCTTTCAAGAGAC 1032

QY 165 -----LEUARGCYLSLEUGLULPHESERALETHRALALASERSEPROGLYPRO 181

DB 1033 ACTCCATTCTCGAGCTGTTCAAC-----ACGGGCCACACCAACATCCCAT 1080

QY 182 GLINLEUARGLEUCYSGINVALSERGLYLEULEUPROLEUARGPROGLYSESERLEUARG 201

DB 1081 AAGGTGCACACCTGTCACACAGAGGTCTGTATCCACCTGGAAACGAGAGATCCAT 1140

QY 202 ILEARGTHRLU-----PROTRPALHISLEULYSALAIPROPHLEUTHRTYR 218

DB 1141 CTGAAGGACATTCACAAAGATCGCAATGCACTTGTCCGGAGGAAACCAACCAAGCTAC 1200

QY 219 PHEGLYLEUPHEGLINVAL 224

DB 1201 TTTGCACTTCAAGGTG 1218

RESULT 14

ABL21473

ID ABL21473 standard; DNA: 978 BP.

XX

XX ABL21473:

AC

XX

XX 26-MAR-2002 (first entry)

XX

XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 15892.

XX

XX Drosophila: developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ds.

XX

XX Drosophila melanogaster.

OS

XX

XX WO200171042-A2.

XX

XX 27-SEP-2001.

XX





50 Sequence 1630 BP: 436 A: 355 C: 379 G: 460 T: 0 other:

## Alignment Scores:

Pred. No.:	2.95	Length:	1630
Score:	106.50	Matches:	67
Percent Similarity:	38.048	Conservative:	38
Best Local Similarity:	24.288	Mismatches:	106
Query Match:	9.178	Indels:	65
DB:	19	Gaps:	13

US-09-245-198a-2 (1-225) x RAV41377 (1-1630)

```

OY 3 SerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGlySer 22
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 78 TCATGATCTTCCTGGCCCTCTGGGCTGGAGCTGGGCCAGCTGCTCGACATCCCTCG 137
OY 23 TrpAlaThrLeuSerAlaGln---GluProSerGlnGluLeuThrAlaGlnSpArg 41
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 138 TTCCTGTACTTCGAGCGAGATGATGATCTACAGA-----ATATCAGAGACAGC 188
OY 42 -----ArgGluProProGluLeuAsnProGlnThr 51
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 189 ACTCACTGCTTTATAGATCTCGAGACTCCATGAAGAGATTTGCAGAGCTCGACT 248
OY 52 GluGluSerGlnAspValValPro----- 59
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 249 CTGGAGAGTGAAGACACACTACCTGCTGCTCGACAGAGAGTGAAGACCTTCAGGG 308
OY 60 -----PheLeuGluGlnLeuValArgProArgArg-----SerAlaProLys 73
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 309 GCCGTGAGAGAGACTGCAACACATGTGGGGCCACAGCGCTTCTCAGAGACTCCAGCT 368
OY 74 -----GlyArgLysAlaArgProArgArgAlaIleAla 84
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 369 ATGATGAGAGGCTCATGTTGGATGTGGCCAGCGAGGCCAGGCCAGCCATTT 428
OY 85 AlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGly 104
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 429 GCACACCTCACCACATCAATGCTGCCAGCATCCCATCGGTTCCAT-----AAAGTC 479
OY 105 ThrValSerGlyTyrGluGln-----ThrLysIleAsnSerSerProLeu 120
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 480 ACTCTGCTCTGTGTACACGATCGAGGAGCTGGGCCAGATCTTAACATGACG----- 533
OY 121 ArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr 140
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 534 -----TTAAGCAACGGAACCTAAGGTTAACCAGATGGCTTCTATTACCTGTAC 584
OY 141 -----CysGlnValHisPheAspGluGlyLysAlaVal-----TyrLeuLys 154
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 585 GCCAATCATTTCTTCGTCATCATGAACATCGGAGAGCGTACCTACAGACTATCTTCAG 644
OY 155 LeuAspLeuLeuVal-----AsnGlyValLeuAlaLeuArgCysLeuGluGln 170
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 645 CTGATGCTGTATGCTGTTAAACAGCATCAAAATCCCAATTCATTAACCTGATGAAA 704
OY 171 PheSerAlaThrAlaAlaSerSerProGlyProGlnLeuArgLeu-----CysGlnVal 188
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 705 GAGAGGAGCAGAGAAACCTGCTCGCAATTCGAATTCACCTTTATTCATTAATGTT 764
OY 189 SerGlyLeuLeuProLeuArgProGlySerSerLeuArgIleArgThrLeuProTyrAla 208
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 765 GGGGGAATTTTCAGAGCTCCGAGCTGTGAGAGAAATTTAGCATTCAGGTGCCAACCTTCC 824
OY 209 HisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 224
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 825 CTGCTGATCCGATCAAGATGCGAGCTTCTTGCGCTTCAAGTT 872

```

Search completed: May 8, 2003, 03:42:39  
Job time: 171.65 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 8, 2003, 02:22:12 ; Search time 1765.08 seconds

(without alignments)  
3709.824 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162  
Sequence: 1 VSLGLALACGLLLVWSL.....PWAHLKAPFLYFGILFQVH 225

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODE=frame\_plus\_p2n.model -DEV=tlp  
-Q=/cgn2.1/USPTO.spool/US09245198/runat.06052003\_170239\_9615/app-query.fasta.1.846  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rgc -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdl -LIST=45  
-DOCCALIGN=200 -THR.SCORE=pt -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09245198.qcgn.1.1.3073 -runat.06052003\_170239\_9615 -NCPV=6 -ICPV=3  
-NO\_XLPPY -NO\_MAP -LARGEDEV -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: gb\_ba:\*  
2: gb\_hgt:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_of:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_hgt\_hum:\*  
31: em\_hgt\_inv:\*  
32: em\_hgt\_other:\*  
33: em\_hgt\_mus:\*  
34: em\_hgt\_pin:\*  
35: em\_hgt\_rod:\*  
36: em\_hgt\_mam:\*  
37: em\_hgt\_vrt:\*  
38: em\_sy:\*  
39: em\_hgt\_hum:\*  
40: em\_hgt\_mus:\*  
41: em\_hgt\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1162	100.0	1168	10	AF030100	AF030100 Mus muscu
2	1020	87.8	1236	6	AR140407	AR140407 Sequence
3	1020	87.8	1306	9	AF030099	AF030099 Homo sapi
4	1020	87.8	1353	6	AX201324	AX201324 Sequence
5	1020	87.8	1368	9	AF055872	AF055872 Homo sapi
6	951	81.8	1651	6	AX180714	AX180714 Sequence
7	945.5	81.4	1651	9	BC019047	BC019047 Homo sapi
8	602.5	51.9	203083	2	AC069459	AC069459 Mus muscu
9	602.5	51.9	234182	10	AF063707	AF063707 Mouse DNA
10	549.5	47.3	177703	2	AC016876	AC016876 Homo sapi
11	549.5	47.3	215795	2	AC127470	AC127470 Pan trogl
12	545	46.9	177555	2	AC130192	AC130192 Sus scrof
13	544.5	46.9	161428	2	AC126925	AC126925 Canis fam
14	536.5	46.2	138792	2	AC119115	AC119115 Rattus no
15	503	43.3	153553	2	AC126921	AC126921 Bos tauri
16	425.5	36.3	184006	2	AC098923	AC098923 Rattus no
17	207	17.8	203281	2	AC126237	AC126237 Canis fam
18	113.5	9.8	1656	3	AB073865	AB073865 Drosophila
19	110.5	9.5	1221	3	AF511923	AF511923 Drosophila
20	110.5	9.5	2159	3	AF521176	AF521176 Gallus ga
21	109.5	9.4	892	5	GGA24345	AL353138 Human DNA
22	109.5	9.3	108967	9	AL353138	AL353138 Streptomy
23	108.5	9.3	42210	1	SC1C2	AE004602 Pseudomon
24	107.5	9.3	13243	1	AE004602	AC061974 Homo sapi
25	107	9.2	178262	2	AC061974	AC060789 Homo sapi
26	107	9.2	179383	2	AC060789	AR156433 Sequence
27	106.5	9.2	1630	6	AR156433	AR164147 Sequence
28	106.5	9.2	1630	6	AR164147	AX147987 Sequence
29	106.5	9.2	1630	6	AX147987	AX451897 Sequence
30	106.5	9.2	1694	6	AX451897	AF019048 Mus muscu
31	106.5	9.2	2225	10	AF019048	AX200995 Sequence
32	106.5	9.2	4034	6	AX200995	AX267730 Sequence
33	106.5	9.2	4034	6	AX267730	AX200993 Sequence
34	106.5	9.2	4894	6	AX267728	AX267728 Sequence
35	106.5	9.2	4894	6	AX267728	AX200996 Sequence
36	106.5	9.2	6976	6	AX200996	AX267731 Sequence
37	106.5	9.2	6976	6	AX267731	AC096533 Homo sapi
38	106.5	9.2	157988	2	AC096533	AC090975 Mus muscu
39	106	9.1	177485	2	AC090975	AC073779 Mus muscu
40	106	9.1	207585	2	AC073779	AC073801 Mus muscu
41	106	9.1	222037	2	AC073801	AY070219 Homo sapi
42	105.5	9.1	4412	9	AY070219	L11016 Homo sapien
43	105.5	9.1	6305	9	HUMLYTOXBB	AX472617 Sequence
44	105.5	9.1	24526	6	AX472617	AL451008 Homo sapi
45	105.5	9.1	62370	2	AL451008	

RESULT 1

## ALIGNMENTS

AF030100 1168 bp mRNA linear ROD 20-DEC-1997  
 LOCUS AF030100 Mus musculus TWEAK mRNA, partial cds.  
 DEFINITION AF030100  
 ACCESSION AF030100  
 VERSION AF030100.1 GI:2707220  
 KEYWORDS  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 1168)  
 Hession,C., Garcia,I. and Browning,J.L.  
 Chicheportriche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H., TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis  
 J. Biol. Chem. 272 (51), 32401-32410 (1997)  
 MEDLINE 98070415  
 PUBMED 9405449  
 REFERENCE 2 (bases 1 to 1168)  
 Chicheportriche,Y., Bixler,S., Tizard,R. and Browning,J.  
 REFERENCE Direct Submission  
 TITLE Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center, Cambridge, MA 02142, USA  
 JOURNAL  
 FEATURES  
 source Location/Qualifiers  
 1..1168  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /cell\_type="peritoneal macrophages"  
 <1..679  
 /note="ligand in the TNF family; secreted protein"  
 /codon\_start=2  
 /product="TWEAK"  
 /protein\_id="AAC3517.1"  
 /db\_xref="GI:2707221"  
 /translation="VLSGLALACGLLVVYSLGSNATLSAQBPSOBEFLAEDREP  
 PELNPOTESODVPELEOLVPRPSAKGRARLAAHYEHPRGDGAQAGV  
 DGVSGMERTKINSSPLRYRDIQGEFTVIRAGLYLVCQVHFDGKAVYLLDLVY  
 GVALRCLREFSATRASSPQPLRQVSGLLPRSSLRIRLPMALNAAPLTY  
 FGLFQVH"  
 CDS  
 BASE COUNT 242 a 360 c 298 g 268 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,47e-79 Length: 1168  
 Score: 1162.00 Matches: 225  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 10 Gaps: 0  
 US-09-245-198a-2 (1-225) x AF030100 (1-1168)  
 QY 1 ValLeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuValValSerLeu 20  
 DB 2 GTGCTGAGCCCTGGCGCTGGCGCTGCTGGCGCTGCTGGCGCTGCTGGCGCTG 61  
 QY 21 GlySerTrrPalThrLeuSerAlaGlnGluProSerGlnGluGluLeuThrAlaGln 40  
 DB 62 GGGAGCTGGGCAACGCTGCTGCTGCCAGGAGCTTCTCAGAGAGAGCTGACAGAGAG 121  
 QY 41 ArgArgGluProGluLeuAsnProGlnInrGluGlnGluSerGlnAspValAlaPro 60  
 DB 122 CCGCGGGAGCCCCCTGTAATGATCCCGACAGAGAAAGCAGAGATGTGGACCTTTC 181  
 QY 61 LeuGlnGlnLeuValArgProArgSerAlaProLysGlyArgLysAlaArgProArg 80  
 DB 182 TTGGAAACACTAGTCCGGCTCGAAGCACTGCTCTAAAGCCGCGAAGCCGCGCTCC 241  
 QY 81 ArgAlaIleAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGln 100  
 DB 242 CGAGTATGTCAGACCATATAGATTCACTCGGCCAGAGACAGATGAGAGACAGACA 301  
 QY 101 GlyValAspGlyThrValSerGlyTrrGluGlnThrLysIleAsnSerSerProLeu 120

|||||  
 DB 302 GGTGTGATGGGAGACATGAGTGGCTGGGAGAGACCAAAATCAACAGCTCCAGCCCTCG 361  
 QY 121 ArgTyrAspArgIleIleGlyGluPheThrValIleArgAlaGlyLeuTyrTrrLeuTyr 140  
 DB 362 CCGTACGACCGCAGATTTGGGAAATTTACAGTCATCAGGGCTGGGCTGACTACCTGAC 421  
 QY 141 CysGlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeuValAsn 160  
 DB 422 TTGCAAGTCACTTGTATGAGGAAAGCGTGTCTACCTGACACTGACTGTGTGGTGAAC 481  
 QY 161 GlyValLeuAlaLeuArgCysLeuGlnGluLysSerAlaThrAlaAlaSerProGly 180  
 DB 482 GGTGTCTGGCCCTGGCGCTGCTGGAAATTTCTCAGCCACAGACAGCAAGCTCTCGGG 541  
 QY 181 ProGlnLeuArgLeuGlnValSerGlyLeuLeuProLeuArgProGlySerLeu 200  
 DB 542 CCCAGCTCCGTTTGTGGCAGGCTGTGGCTGTGGCTGGCGGCGGCGGCTTCCCTT 601  
 QY 201 ArgLeuArgThrLeuProTrrPalHisLeuLysAlaAlaProPheLeuThrTrrPheGly 220  
 DB 602 CGGATCGGACACCTCCCTGGGCTCATCTTAAGGCTGGCCCTTCTTACTACTTGA 661  
 QY 221 LeuPheGlnValHis 225  
 DB 662 CTCTTCAAGTTCAC 676  
 RESULT 2  
 LOCUS AR140407 1236 bp DNA linear PAT 16-JUN-2001  
 DEFINITION Sequence 1 from patent US 6207642.  
 ACCESSION AR140407  
 VERSION AR140407.1 GI:14482903  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1236)  
 WILEY,S.R.  
 TITLE Member of the TNF family useful for treatment and diagnosis of disease  
 JOURNAL Patent: US 6207642-A 1 27-MAR-2001;  
 FEATURES  
 source Location/Qualifiers  
 1..1236  
 /organism="unknown"  
 BASE COUNT 225 a 416 c 358 g 237 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 9.38e-69 Length: 1236  
 Score: 1020.00 Matches: 199  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 88.84% Mismatches: 16  
 Query Match: 87.78% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-245-198a-2 (1-225) x AR140407 (1-1236)  
 QY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuValValSerLeuGly 21  
 DB 76 CTGGGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGG 135  
 QY 22 SerTrrPalThrLeuSerAlaGlnGluProSerGlnGluGluLeuThrAlaGln 41  
 DB 136 AGCGGGGATCGCTGTCCGCCAGAGAGCTGCGCAGAGAGAGCTGGTGGTCAAGAGAG 195  
 QY 42 ArgGluProGluLeuAsnProGlnInrGluGlnGluSerGlnAspValAlaProPhe 61  
 DB 196 CAGGACCCCTCGCACTGATCCCGACAGAGAAAGCAGAGATCTGCGCTTCCG 255  
 QY 62 GlnGlnLeuValArgProArgSerAlaProLysGlyArgLysAlaArgProArg 81  
 DB 256 AACGACTATGTCGCTCGAAGAGTGCACCTAAAGCGCGGAAACACGCGCTCGAGA 315

```

QY      82  AAlaIleAlaAlaHsTYrGluValAlHisProArpProGlyGlnAspGlyAlaGlnAlaGly 101
      |||
Db      316  GCGATCCGAGCCCATTTATGAAATTCATCCAGCAGTGGACAGGAGGAGCCAGCAGAGT 375
QY      102  ValAspGlyThrValSerGlyTYrPgluGluThrLysIleAsnSerSerProLeuArg 121
      |||
Db      376  GTGAGCGGAGACAGTACGAGTGGAGAGAACAGAAATCAACAGCTCCAGCCCTGTGGCC 435
QY      122  TYrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTYrTYrLeuTYrCys 141
      |||
Db      436  TACACACGCGCAGATCGGGAGATTATAGTACACCGCGGTGGCTCTACTACTACTACTGT 495
QY      142  GlnValHisPheAspGluGlyLysAlaValTYrLeuLysIleAsnSerProLeuValAsnGly 161
      |||
Db      496  CAGGTGCACCTTGATGAGAGGAGAGGCTGTCTACTGAAAGCTGAGCTGTGTGTGTGTGTGT 555
QY      162  ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyPro 181
      |||
Db      556  GTGCTGGCCCTGCGCTGCGCTGGAGAAATTCACAGCCAGTGGCGAGTTCCTCGGGGCC 615
QY      182  GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
      |||
Db      616  CAGCTCCGCTCTGCGCAGGTGTGTGGCTGTGGCCCTGCGCGCAGGCTCTCTCTCTCTCTCT 675
QY      202  IleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTYrPheGlyLeu 221
      |||
Db      676  ATCCGACCCCTCCCTGGGCCCATCTCAAGCTGCCCCCTTCTCCTACTCTGAGACTC 735
QY      222  PheGlnValHis 225
      |||
Db      736  TTCAGGTTTCAC 747

RESULT 3
AF030099 1306 bp mRNA linear PRI 20-DEC-1997
DEFINITION Homo sapiens TWEAK mRNA, complete cds.
ACCESSION AF030099
VERSION AF030099.1 GI:2707218
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
            Homo sapiens
            Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1306)
            Chicheportiche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.,
            Hession,C., Garcia,I. and Browning,J.L.
            TWEAK, a new secreted ligand in the tumor necrosis factor family
            that weakly induces apoptosis
            J. Biol. Chem. 272 (51), 32401-32410 (1997)
JOURNAL MEDLINE 9405449
TITLE
JOURNAL 2 (bases 1 to 1306)
AUTHORS Bourdon,P., Hession,C., Tizard,R. and Browning,J.
TITLE Direct Submission
JOURNAL Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,
Cambridge, MA 02142, USA
FEATURES
SOURCE
1. 1306
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /chromosome="17"
    /map="17p13"
    /class_type="tcons11"
    /note="Ligand in the TNF family; secreted protein; start
    codon not verified experimentally."
    /codon_start=1
    /product="TWEAK"
    /protein_id="AAC51923.1"
    /db_xref="GI:2707219"
    /translacion="MAARSRQRRGRGEPGALVPLALGLIALAGLILAAVSL
    GRSASLSAQEPADQELVAEDQPSLEINPQTESQDPAPLNLVRRRSAPGRKTR
  
```

```

ARRAIAHYEHPRPGDGAQAGVDGVSQMEPARINSSPLRYNRIGEIFYTRAGL
YLYGVHPDEGKAVYLIKLDLYDVYALRCLAEFEFSATASSLSGLPOLRQVSGLLAL
RPSSLRIRLPAHLKAPFLRYRFLRYVH"
BASE COUNT 247 a 434 c 368 g 257 t
ORIGIN
Alignment Scores:
Pred. No.: 9,95e-69 length: 1306
Score: 1020.00 Matches: 199
Percent Similarity: 92.86% Conservative: 9
Best Local Similarity: 88.84% Mismatches: 16
Query Match: 87.78% Indels: 0
DB: 9 Caps: 0
US-09-245-198a-2 (1-225) x AF030099 (1-1306)
QY      2  LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21
      |||
Db      93  CTGGGCGTGGGCTGGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGG 152
QY      22  SerThrAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGluAspArg 41
      |||
Db      153  AGCGGGGCGATGCTGTCCGCCAGAGAGCTCCAGAGAGAGTGTGTGGCAGAGAGAGAC 212
QY      42  ArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPheLeu 61
      |||
Db      213  CAGAGCCCGTGGAGACTGAATCCACAGACAGAAACCCAGATCCCTGGCCCTTCTCTG 272
QY      62  GlnGlnLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArgProArgArg 81
      |||
Db      273  AACCGACTAGTTCGGCTCCGAGAAAGTGCACCTAAAGCCGGAACACGGGCTCGAAGA 332
QY      82  AAlaIleAlaAlaHsTYrGluValAlHisProArpProGlyGlnAspGlyAlaGlnAlaGly 101
      |||
Db      333  GCGATCCGAGCCCATTTATGAAATTCATCCAGCAGTGGACAGGAGGAGCCAGCAGAGT 392
QY      102  ValAspGlyThrValSerGlyTYrPgluGluThrLysIleAsnSerSerProLeuArg 121
      |||
Db      393  GTGAGCGGAGACAGTACGAGTGGAGAGAACCCAGAAATCAACAGCTCCAGCCCTGTGGCC 452
QY      122  TYrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTYrTYrLeuTYrCys 141
      |||
Db      453  TACACACGCGCAGATCGGGAGATTATAGTACACCGGCTGGCTCTACTACTACTACTGT 512
QY      142  GlnValHisPheAspGluGlyLysAlaValTYrLeuLysIleAsnSerProLeuValAsnGly 161
      |||
Db      513  CAGGTGCACCTTGATGAGGAGGAGGAGCTGTCTACTGAAAGCTGAGCTGTGTGTGTGTGT 572
QY      162  ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyPro 181
      |||
Db      573  GTGCTGGCCCTGCGCTGCGCTGGAGAAATTCACAGCCAGTGGCGCAGTTCCTCGGGGCC 632
QY      182  GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
      |||
Db      633  CAGCTCCGCTCTGCGCAGGTGTGTGGCTGTGGCCCTGCGCGCAGGCTCTCTCTCTCTCT 692
QY      202  IleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTYrPheGlyLeu 221
      |||
Db      693  ATCCGACCCCTCCCTGGGCCCATCTCAAGCTGCCCCCTTCTCCTACTCTTGGACTC 752
QY      222  PheGlnValHis 225
      |||
Db      753  TTCAGGTTTCAC 764

RESULT 4
AX201324 1353 bp DNA linear PAT 30-AUG-2001
LOCUS AX201324
DEFINITION Sequence 3 from Patent WO0153486.
ACCESSION AX201324
VERSION AX201324.1 GI:15391154
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
  
```



```

QY 62 GlnGlnLeuValArgProArgSerAlaProLysGlyArgLysAlaArgProArgArg 81
Db 313 AACCGACTAGTGGCGCTGCAGAACTGCACCTAAAGCGCGAAGAACAGCGGCTCCAGAA 372
QY 82 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
Db 373 GCGATCGCAGCCCATATATAGTTCATCCAGCAGCTGAGACAGACGAGCGCAGCAGGT 432
QY 102 ValAspArgLysThrValSerLysLysLysLysLysLysLysLysLysLysLysLys 121
Db 433 GTGGACGGACAGATGAGTGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 492
QY 122 TyrAspArgGlnIleGlyLysLysLysLysLysLysLysLysLysLysLysLysLys 141
Db 493 TACACCGCGACATCGCGGAGTATATAGTTCACCCGCGGCTGCTACTACTACTACTACT 552
QY 142 GlnValHisPheAspGlnGlyLysLysLysLysLysLysLysLysLysLysLysLys 161
Db 553 CAGGTGCACTTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 612
QY 162 ValLeuAlaAlaLeuArgCysLeuGlnGlnLysLysLysLysLysLysLysLysLys 181
Db 613 GTGCTGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 672
QY 182 GlnLeuArgLeuGlnGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
Db 673 CAGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 732
QY 202 IleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221
Db 733 ATCCGACCCCTCCCTGGGCGCATCTCAAGGCTGCGCCCTCTCTCACTACTGCGAGTC 792
QY 222 PheGlnValHis 225
Db 793 TTCAGAGTTCAC 804

RESULT 6
LOCUS AX180714 898 bp DNA linear PAT 06-AUG-2001
DEFINITION Sequence 1 from Patent WO0145730.
ACCESSION AX180714
VERSION AX180714.1 GI:15132570
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source location/Qualifiers
1..898
/organism="synthetic construct"
/db_xref="taxon:32630"
52..873
/note="human TWEAK fusion protein construct"
/codon_start=1
/transl_table=11
/protein_id="CAC50004.1"
/db_xref="GI:15132571"
/translation="MATGRTSLIAFGILCPMLQESGATSSDDMKQIEDKIEEILS
KIYHENEIARIKKIKIGERTSSISGASISAOEBOEIVAEQDQSEINPQIEES
ODPAFLNRLVPRRSAPKGRKTRARRAIAHAYEHVPRGDDAGAGVGYSGWEAA
KINSSEPLRNQIEFTVTRAGLYLVCQVHFDEGKAVYLKLDLVLGCVIALRGLEE
FSATPAASLIGDRLRCQVSGLALRPGSSLRIRLTPMAHLKAPLTYFGLEQVH"
BASE COUNT 187 a 266 c 267 g 178 t
ORIGIN

```

```

Best Local Similarity: 88.89% Mismatches: 14
Query Match: 81.84% Indels: 0
DB: 6 Gaps: 0

US-09-245-198a-2 (1-225) x AX180714 (1-898)
QY 19 SerLeuGlySerThrAlaThrLeuSerAlaGlnGlnProSerGlnGlnLeuThrAla 38
Db 250 AGTTTGGGAGCCCGGCGATCGCTGCTGCCCGACAGACCTGCGCCAGAGAGAGCTGGTGGCA 309
QY 39 GluAspArgGlnGluProGlnLeuAsnProGlnThrGlnGlnSerGlnAspValVal 58
Db 310 GAGGACAGCAGAGACCCGTCGGAAGTATCCCAACAGAGAAAGAACAGAGTCTCTGCG 369
QY 59 ProPheLeuGlnGlnLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArg 78
Db 370 CCTTTCCTACACCGACTATGCTGGCTCCAGAGGACCATTAAGGCGCGGAAACACCG 429
QY 79 ProArgArgAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGlyAla 98
Db 430 GCTCGAGAGCCGATCCGACCCATTTGAAATTATCCAGACTGAGACAGAGGAGCGG 489
QY 99 GlnAlaGlyValAspGlyThrValSerGlyTyrGlnGlnLysLysAsnSerSer 118
Db 490 CAGGACAGTGTGGACGAGCAGTGAAGTGGCTGGAGAGAACCAAGATCAACAGCTCCAGC 549
QY 119 ProLeuArgTyrAspArgGlnIleGlyLysLysLysLysLysLysLysLysLysLys 138
Db 550 CCTTGGCTGACACACCGCAGATCGGGGAGTTATAGTACCCGCGGCTGCTACTACTAC 609
QY 139 LeuTyrCysGlnValHisPheAspGlnGlyLysLysLysLysLysLysLysLysLys 158
Db 610 CTGTACTGTCAGAGTGCATTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 669
QY 159 ValAsnGlyValLeuAlaLeuArgCysLeuGlnGlnLysLysLysLysLysLysLys 178
Db 670 GTGATGATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 729
QY 179 ProGlyProGlnLeuArgCysGlnValSerGlyLeuLeuProLeuArgProGlySer 198
Db 730 CTCGGGCGCCAGCTCCGCTGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 789
QY 199 SerLeuArgIleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThr 218
Db 790 TCCCTCGGATCCGACCCCTCCCTGCGCCATCTCAAGGCTGCGCCCTCTCTCACTAC 849
QY 219 PheGlyLeuPheGlnValHis 225
Db 850 TTCGACTCTTCAGAGTTCAC 870

RESULT 7
LOCUS BC019047 1651 bp mRNA linear PRI 11-DEC-2001
DEFINITION Homo sapiens, similar to tumor necrosis factor (ligand)
superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA,
complete cds.
ACCESSION BC019047
VERSION BC019047.1 GI:17512138
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
USA NIH-MGC Project URL: http://mgc.mcl.nih.gov
Contact: MGC help desk
Email: cgapbs-rt@mail.nih.gov

```







FEATURES This sequence is the entire insert of clone RP23-422L16.  
Location/Qualifiers

1. 234182

/organism="Mus musculus"

/db\_xref="taxon:10090"

/chromosome="11"

/clone="RP23-422L16"

/clone\_11b="RP23-23"

/complement(84050..84311)

/note="Sequence from uni-directional primer reads and dGTP

big dye terminator reads only."

BASE COUNT 59310 a 56824 c 57519 g 60529 t

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-245-198a-2 (1-225) x AL603707 (1-234182)

QY 100

DB 66609

QY 120

DB 66549

QY 140

DB 66489

QY 141

DB 66429

QY 142

DB 66369

QY 146

DB 66309

QY 166

DB 66249

QY 186

DB 66189

QY 206

DB 66129

QY 225

DB 66072

QY 225

DB 66072

QY 225

DB 66072

QY 225

DB 66072

QY 225

DB 66072

QY 225

DB 66072

QY 225

DB 66072

QY 225

DB 66072

QY 225

AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Bliren, B., Nusbaum, C., and Lande, E.  
Homo sapiens chromosome, clone RP11-186B7  
Unpublished  
2 (bases 1 to 177703)  
Bliren, B., Linton, L., Nusbaum, C., Lande, E., Allen, N., Anderson, M.,  
Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhgalter, B.,  
Brown, A., Castle, A., Colangelo, M., Collins, S., Collamore, A.,  
Cooke, P., Deatellano, K., Dewar, K., Domini, M., Donelan, L., Doyle, M.,  
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,  
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heatford, A., Horton, L.,  
Howard, J. C., Johnson, R., Jones, C., Kenn, L., Karatas, A., Klein, J.,  
Lehoczky, J., Liu, C., Locke, K., Macdonald, P., Margis, N.,  
McEwan, P., McGuire, A., McKernan, K., McLaughlin, J., Meltrini, J.,  
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Testfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
Wyman, D., Ye, M. J., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Direct Submission  
Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 177703)  
Bliren, B., Nusbaum, C., Lande, E., Allen, N., Anderson, S.,  
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,  
Camarata, J., Chang, J., Chazaro, B., Choepel, X., Collamore, A.,  
Cooke, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J. S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,  
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamt, A.,  
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Punkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,  
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
Testfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT

Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 2, 2002 this sequence version replaced g1:15421989.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www.seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
Project Information  
Center project name: L3849  
Center clone name: 186\_B\_7

NOTE: This is a 'working draft' sequence. It currently  
consists of 10 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 48645: contig of 48645 bp in length  
\* 48646 48745: gap of 100 bp  
\* 48746 48745: gap of 100 bp  
\* 58278 58377: contig of 9532 bp in length  
\* 58278 58377: gap of 100 bp  
\* 58378 58377: contig of 7427 bp in length  
\* 58378 58377: gap of 100 bp  
\* 65805 65904: contig of 100 bp  
\* 65905 65904: gap of 100 bp  
\* 79793 79793: contig of 13889 bp in length  
\* 79794 79893: gap of 100 bp  
\* 79894 79893: contig of 19600 bp in length









```

Db 42506 GCGGCTCTGCGCCGCCCGCCAGTGCACCTTGTATGAGGGAGGAGCTGTCTACTAGCT 42447
Oy 155 uaspleuauValaanglyValleuAlaueuArgCysleugluglupheserAlaThrAl 175
Db 42446 GCACTTGGTGGTGGAGACGCCCTGGCCCTGGCTGGCCGAGAGGTCTTCGCCACAC 42387
Oy 175 aAlaserSerProglIyProglInleuArgleuCysglInValSerGIyleuLeuProleuAr 195
Db 42386 CGCCAGCAGCGCTGGGCCGCCAGCTCGCTCGCCAGAGTGTGGGCTTTCCCTCCG 42327
Oy 195 gProglIySerSerleuArgIleArgThrleuProTrpAlaHisleuValaAlaProPh 215
Db 42326 GCCCGGGTCTCTCCCTCGCGGATCCGACCCCTCGGCGCCATCAGGCGGCCCTT 42267
Oy 215 eleuthTyrrPheglIyleuPheglInValHis 225
Db 42266 CCTTACTACTCTGGAGCTCTTCCAGGTAC 42236

```

```

RESULT 14
AC119115
LOCUS
DEFINITION Rattus norvegicus clone CH230-320N23, *** SEQUENCING IN PROGRESS
ACCESSION AC119115
VERSION AC119115.2 GI:21746718
KEYWORDS HTG; HTGS; PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

## REFERENCE

## AUTHORS

1 (bases 1 to 138792)  
 Munzy D.M., Adams C., Adio-Oduola B., Ali-oshan F.R., Allen C.,  
 Aishbrooks S.L., Amaratunga H.C., Are J.R., Ayalew M., Banks T.,  
 Barabak J., Benton J., Bimaga K., Blankenburg K., Bonnah D.,  
 Bouck J., Bowie S., Brileva M., Brown E., Brown M., Bryant N.P.,  
 Buhay C., Burch P., Burkett C., Burrell K.L., Byrd N.C.,  
 Caron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D.,  
 Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C.,  
 Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,  
 Davila M.L., Davis C., Davy-Carroll L., Dederich D.A.,  
 Delaney R.K., Delgado O., Denn A.L., Ding Y., Dinh H.H.,  
 Doutheaite K.J., Draper H., Dugan-Rocha S., Durbin K.J.,  
 Earnhart C., Edgar D., Edwards C.C., Elhaj C., Escotto M.,  
 Falls T., Ferraguto D., Flagg N., Ford J., Foster P., Frantz P.,  
 Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R.,  
 Gorrell U.H., Guevara W., Gunaratne P., Hale S., Hamilton K.,  
 Harris C., Harris R., Hart M., Havlak P., Hawes A., Hernandez J.,  
 Hernandez O., Hodgson A., Hognes M., Hollway C., Hollins B.,  
 Homsl F., Howard S., Huber J., Hulv S., Hume J., Jackson L.E.,  
 Jacobson B., Jia Y., Johnson R., Jolivet S., Joudh S.,  
 Karlsson E., Kelly S., Khan U., King L., Korvah U., Kovar C.,  
 Kralovic J., Kuresh A., Landry N., Leal B., Lewis L.C., Lewis L.,  
 Li J., Li Z., Licharge O., Lieu C., Liu J., Liu W., Loulseged H.,  
 Lozad R.J., Lu X., Lucier A., Lucier R., Luna R., Ma J.,  
 Maheshwari M., Mapua P., Martin R., Martindale A., Martinez E.,  
 Massey E., Mawhney E., McLeod M.P., Meador M., Mei G., Metzger M.,  
 Miner G., Miner Z., Mitchell T., Monabaty K., Morgan M., Morris S.,  
 Moser M., Neal D., Newton J., Newton N., Nguyen A., Nguyen N.,  
 Nguyen N., Nickerson E., Nwokemo S., Oguh M., Okunolu G.,  
 Oragunye N., Oviedo R., Pace A., Payton B., Peery M., Perez L.,  
 Peters L., Pickens R., Primus E., Pu L.L., Quiles M., Ren Y.,  
 Rivers M., Rojia A., Rojibokan I., Rolfe M., Ruiz S., Savery G.,  
 Soedergren E., Sonalke T., Sparks S., Shoshart N., Sisson I.,  
 Sulten A., Svatek A., Tabor P., Tamerisa A., Tamerisa K., Tang H.,  
 Tansey J., Taylor C., Taylor T., Telford B., Thomas N., Thomas S.,  
 Umanu K., Vasquez L., Vera V., Villalón D., Vinson R., Wang O.,  
 Wang S., Ward-Moore S., Warren R., Washington C., Watlington S.,  
 Williams G., Williamson A., Wleczek R., Woodson S., Worley K.,  
 Wu C., Wu Y., Wu F., Zhou J., Zorrilla S., Nelson D.,  
 Weinstock G. and Gibbs R.

## TITLE

Direct Submission  
 Unpublished  
 2 (bases 1 to 138792)  
 Worley K.C.

## REFERENCE

Submitted (25-Apr-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 138792)  
 Worley K.C.

## AUTHORS

Submitted (18-Jul-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 14, 2002 this sequence version replaced gi:20303440.

## COMMENT

Center: Baylor College of Medicine  
 Center code: BCM  
 Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: hgsc-help@bcm.tmc.edu  
 Project Information  
 Center project name: G0XG  
 Center clone name: CH230-320N23  
 Summary Statistics  
 Sequencing vector: Plasmid  
 Chemistry: Dye-terminator Big Dye, 100% of reads  
 Assembly program: Phrap, version 0.990329  
 Consensus quality: 112720 bases at least Q40  
 Consensus quality: 116666 bases at least Q30  
 Consensus quality: 119165 bases at least Q20  
 NOTE: Estimated insert size may differ from sequence length  
 (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 NOTE: This is a working draft sequence. It currently  
 consists of 32 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1	1022:	contig of 1022 bp in length
1023	1122:	gap of unknown length
1123	2394:	contig of 1272 bp in length
2395	2494:	gap of unknown length
2495	3663:	contig of 1169 bp in length
3664	3763:	gap of unknown length
3764	4879:	contig of 1116 bp in length
4880	4979:	gap of unknown length
4980	6787:	contig of 1808 bp in length
6788	6887:	gap of unknown length
6888	8192:	contig of 1305 bp in length
8193	8292:	gap of unknown length
8293	9461:	contig of 1169 bp in length
9462	9561:	gap of unknown length
9562	11435:	contig of 1874 bp in length
11436	11535:	gap of unknown length
11536	13535:	contig of 2000 bp in length
13536	13635:	gap of unknown length
13636	17323:	contig of 3688 bp in length
17324	17424:	gap of unknown length
17424	20406:	contig of 2983 bp in length
20407	20506:	gap of unknown length
20507	23445:	contig of 2939 bp in length
23446	23545:	gap of unknown length
23546	25600:	contig of 2055 bp in length
25601	25700:	gap of unknown length
25701	28049:	contig of 2349 bp in length
28050	28149:	gap of unknown length
28150	30474:	contig of 2325 bp in length
30475	30574:	gap of unknown length
30575	35498:	contig of 4924 bp in length
35499	35598:	gap of unknown length

*	35559	39284:	contig of 3686 bp in length
*	35985	39384:	gap of unknown length
*	39385	41970:	contig of 2586 bp in length
*	41971	42070:	gap of unknown length
*	42071	45653:	contig of 3589 bp in length
*	45660	45759:	gap of unknown length
*	45760	50333:	contig of 4573 bp in length
*	50333	50433:	gap of unknown length
*	50433	54693:	contig of 4263 bp in length
*	54696	54793:	gap of unknown length
*	54796	58283:	contig of 3494 bp in length
*	58290	58389:	gap of unknown length
*	58390	63599:	contig of 5210 bp in length
*	63599	63699:	gap of unknown length
*	63700	70305:	contig of 6606 bp in length
*	70306	70405:	gap of unknown length
*	70406	76123:	contig of 5718 bp in length
*	76124	76223:	gap of unknown length
*	76224	84961:	contig of 8738 bp in length
*	84962	85061:	gap of unknown length
*	85062	93714:	contig of 8553 bp in length
*	93615	93714:	gap of unknown length
*	93715	103352:	contig of 9638 bp in length
*	103353	103452:	gap of unknown length
*	103453	110299:	contig of 6647 bp in length
*	110300	110399:	gap of unknown length
*	110400	116180:	contig of 5781 bp in length
*	116181	125486:	gap of unknown length
*	116581	125486:	contig of 9204 bp in length
*	125485	125588:	gap of unknown length
*	125585	128792:	contig of 13208 bp in length

[illegible]

Alignment Scores:	
Pred. No.:	7, 23e-30
Score:	536.50
Percent Similarity:	67.468
Best Local Similarity:	65.098
Query Match:	46.17%
BB:	2
Gaps:	1
Length:	1387932
Matches:	110
Conservative:	4
Mismatches:	2
Indels:	53

US-09-245-198A-2 (1-225) x AC119115 (1-138792)

QY	101	GLYVALAPGGLYTHRVALSERGLYTRPOLGLVHRLYSILEANSERSETSERPROLEU	120
Db	137345	GGTGTGGATGGGACAGTACGTAGTGGCTGGGACAGACCAAAATCAACACTCCAGCCCTGTG	137400
QY	121	ARGTYRASPARGINILEGLYGLUPHENYRVALILEARGALAGLYLEUITYRYTYRLEUTYR	140
Db	137405	CGCATATGACCCGCCAGATTGGGGAAATTACGGTCATCAAGCGTGGGCTCTACTACCTGTAC	137460
QY	141	CYS-----	141
Db	137465	TGTCAAGTAAGCCCTGGGCTCATGGGTAGAGCGATGGCTAAGGGAGGAAGCGTTGGCA	137520
QY	141	-----	141
QY	141	-----	141
Db	137525	AGGAATGGGTGGGAGTGGGAGAACCTGGTTTCATGAAGAAAGTCCTGGATTTCGATG	137580
QY	142	-----GLNVALIHSPHEASPGLU	147
Db	137585	AGGGCAGGCACAGGCTCTGGATTTCGTGGTTTCTCTGTGCTCCAGGCGCACTTTGATGAG	137640
QY	148	GLYLYSALAVALTYRLEULYSLEUASPLEULEUVALASNGLYVALLEUALALEUARGLYCYS	167
Db	137645	GGGAAAGCAGCTCTACCTGAAAGCTGGAATCTGCTGGTGAATGGATGTGCTGGCCCTGGCGCTGC	137700

QY 168 LeGIuGIuPheSerAlaThrAlaAlaSerSerProGIuProGInLeuAluGIuCyG6In 187  
Db 137705 CTGGAGAGATTCTAGGCACACGACGACGCTCTCTGGGCCCCAGCTCGCTGTGGCAG 137764  
QY 188 ValSerGIuLeuLeuProLeuArgProGIuSerSerLeuArgIleArgTInLeuProTyr 207  
Db 137765 GTGCTGGGGCTGTGGCCCTCTCGACACAGGGCTTCCTCCCTGGAGACGCTACCACTTGC 137824  
QY 208 AlaHisLeuAlaAlaAlaProPheLeu 216  
Db 137825 AGTCATATTAAAGCGCGCTACCTTGTT 137851

RESULT_15	
AC126921	
LOCUS	AC126921
DEFINITION	AC126921 15355 bp DNA linear HTG 10-JUL-2002
ACCESSION	Bos taurus clone RP42-45D24, WORKING DRAFT SEQUENCE, 12 unordered
VERSION	pieces.
KEYWORDS	AC126921
SOURCE	AC126921.1 GI:21724098
ORGANISM	HTG; HTGS_PHASE1; HTGS_DRAFT.
	COW.
	Bos taurus

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 2312: contig of 2312 bp in length  
\* 2313 2412: gap of unknown length



Accession	Contig	Length
2413	contig of 3429 bp in length	
5842	gap of unknown length	
5942	contig of 2494 bp in length	
8436	gap of unknown length	
8536	gap of unknown length	
15800	contig of 7264 bp in length	
15900	gap of unknown length	
25224	contig of 9325 bp in length	
25225	gap of unknown length	
25325	contig of 7180 bp in length	
32505	gap of unknown length	
32605	contig of 8366 bp in length	
40971	gap of unknown length	
41071	contig of 15320 bp in length	
56591	gap of unknown length	
56691	contig of 17079 bp in length	
73770	gap of unknown length	
73870	contig of 16590 bp in length	
90860	gap of unknown length	
90960	contig of 20469 bp in length	
111429	gap of unknown length	
111528	contig of 42025 bp in length	
115529	gap of unknown length	

```

source
1. .13353
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="RP42-45D24"
/clone_id="RP42"
1. .2312
/note="assembly-fragment"
2413. .5841
/note="assembly-fragment
clone_end:T7
vector_side:left"
5942. .8435
/note="assembly-fragment"
8536. .15799
/note="assembly-fragment"
15900. .25224
/note="assembly-fragment"
25325. .32504
/note="assembly-fragment"
32605. .40970
/note="assembly-fragment"
41071. .56590
/note="assembly-fragment"
56691. .73769
/note="assembly-fragment"
73870. .90859
/note="assembly-fragment"
90960. .111428
/note="assembly-fragment"
111529. .153553
/note="assembly-fragment
clone_end:SP6
vector_side:right"
39178 a 33933 c 36616 g 39259 t 1107 others

```

Alignment Scores:	
Pred. No.:	2,81e-27
Score:	503.00
Percent Similarity:	58.59%
Best Local Similarity:	56.06%
Query Match:	43.29%
DB:	2
	Gaps: 1
	length: 153553
	Matches: 111
	Conservative: 5
	Mismatches: 17
	Indels: 66
	Gaps: 1

US-09-245-198A-2 (1-225) x AC126921 (1-153553)

QY	93	ProGlyGlnAspGlyAlaIleAlaValaAspGlyThrValSerGlyTrpGluIuThr	112
	111	::: ::	
	:::	::: ::	
Db	26848	CCTGCTCTGAAACACCACATTCAGGTGTGGACGGGACGGTAGTGGCTGGGACGAGGCC	26907
QY	113	LysIleasnSerSerSerProleuArgTrpAspArgGlnIleGlyGluIuThrValIle	132

Db	26908	AAAAACCAACAGTCCCAACCCCTGGCGTATGACTGCCAGACCCGGGCAATTACGGTCACC	26967
QY	133	ArgAlaGlyLeuTyrTyrLeuTyrCysGlnVal	143
Db	26968	CGGGCTGGGCTGTACTACTGTACTGTCTAGTCAAGT-AAAGCCCACTGAGCTCCAGGGGTAAAG	27026
QY	143	-----	143
Db	27037	CCGGAACGTAAAGAGAGAAAGGCTGGGCTTCGGGGTTGGGGCAAGTTAAAGTGGGAGGG	27086
QY	143	-----	143
Db	27087	GAGCCTGGGGTTGGGCTGAGAGAGACCTTGGGCTCTAAGACACCTGATGAAGCCC	27146
QY	144	-----HisPheaspGlu	147
Db	27147	AGGGCCACAGAGAGGCGCTGAGACTCCGCCCTCCCTGCGCCGCCAGGTGACCTTTGATGAG	27206
QY	148	GlyValAlaValTyrLeuLysLysaspLeuValAsnGlyValLeuAlaLeuArgCys	167
Db	27207	GGGAAGGCTGTCTACTGAAGCTGACCTTGCTGGTGAGACACCGCTGGCCCTGGGCTGC	27266
QY	168	LeuGlnGluPheSerAlaThrAlaAlaSerProGlyProGlnLeuArgLeuCysGln	187
Db	27267	CTGGAGAGAAATTCTGGGCCACTCGGCCCACTTCCCTGGGCCCACTCCGTCTGTGCCAA	27326
QY	188	ValSerGlyLeuLeuProLeuArgProGlySerSerLeuArgIleArgThrLeuProTrp	207
Db	27327	GTGTAGGGGCTCTTCCGCTCGCGGAGAGGTTCTCCCTCGGATCCGACCCCTCCGCCAG	27386
QY	208	AlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis	225
Db	27387	ACCCAACTCAAGGGCTGCCCTCTCTCACTTACTTGGACTCTTCCAAAGTTTAC	27440

Search completed: May 8, 2003, 03:40:19  
Job time : 2022.08 secs

40

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:07:02 ; Search time 28.7328 Seconds

(without alignments)  
1043.455 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162

Sequence: 1 VLSTGLTACTGLLVVSL.....PMAHLKAPFLTYGLRQVH 225

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1162	100.0	225	19	AAW47524
2	1162	100.0	225	21	AAW47524
3	1089	93.7	225	21	AAW47524
4	1020	87.8	249	19	AAW29745
5	1020	87.8	249	20	AAW29745
6	1020	87.8	249	21	AAW29745
7	1020	87.8	249	21	AAW29745
8	1020	87.8	249	22	AAW29745
9	1020	87.8	249	23	AAW29745
10	1020	87.8	284	19	AAW47525

11	951	81.8	208	20	AAW93590	Human TNFR3 protel
12	951	81.8	273	22	AAW03499	TWEAK extracellular
13	742	63.9	189	19	AAW29746	TNF related endothe
14	742	63.9	189	22	AAW08952	Human UTA1lag TREP
15	721	62.0	146	22	AAE00895	Human TREPA (TNF r
16	111	9.6	406	23	AAW77717	Drosophila melanog
17	110.5	9.5	409	23	AAW77718	Drosophila melanog
18	107.5	9.3	325	22	ABW67553	Drosophila melanog
19	106.5	9.2	294	19	AAW69956	Mouse OBW protein
20	106.5	9.2	294	19	AAW68232	NF-kB receptor act
21	106.5	9.2	294	19	AAW68737	NF-kB receptor act
22	106.5	9.2	294	22	AAE04455	Murine receptor ac
23	106.5	9.2	294	22	AAE01992	Murine RANKL (rece
24	103	8.9	220	22	AAW62340	Gp120 v3 loop-CD15
25	102.5	8.8	316	19	AAW83194	Human osteoprotege
26	102.5	8.8	316	19	AAW83017	Osteoclastogenesis
27	102.5	8.8	316	19	AAW59654	Amino acid sequenc
28	102.5	8.8	316	20	AAW17874	Murine TRANCE. Mu
29	102.5	8.8	316	21	AAW91024	Mouse OBW protein
30	102.5	8.8	316	21	AAW84418	Amino acid sequenc
31	102.5	8.8	316	21	AAW84419	Amino acid sequenc
32	102.5	8.8	316	22	AAW8289	Mouse TRANCE prote
33	99	8.5	234	22	AAW62339	Gp120 v3 loop-CD15
34	97	8.3	255	22	AAW66311	Fusion construct C
35	96.5	8.3	409	23	AAW77716	Drosophila melanog
36	92	7.9	261	23	AAW49225	Human CD145 mutant
37	91	7.8	182	18	AAW09128	CD40 ligand/zipper
38	91	7.8	182	18	AAW09123	CD40 ligand/zipper
39	90.5	7.8	211	21	AAW58216	Canine mature CD15
40	90.5	7.8	260	21	AAW58215	Canine CD154. Can
41	90.5	7.8	318	22	AAW82092	Rat osteoclast dif
42	90.5	7.8	531	10	AAW91776	Germ cell alkaline
43	90	7.7	182	18	AAW24011	CD40 ligand/zipper
44	90	7.7	182	18	AAW24013	CD40 ligand/zipper
45	90	7.7	261	18	AAW09115	Human CD40L muteln

#### ALIGNMENTS

RESULT 1	
AAW47524	AAW47524 standard; Protein: 225 AA.
ID	AAW47524 standard; Protein: 225 AA.
AC	AAW47524;
XX	
DT	21-JUL-1998 (first entry)
XX	
DE	Mus musculus tumour necrosis factor related ligand (TRELL).
XX	
KW	TRELL: tumour necrosis factor related ligand; tnfr; treatment;
KW	cancer; autoimmune disease; immune system; stimulation; suppression;
KW	graft rejection.
XX	
OS	Mus musculus.
XX	
FT	Key
FT	Domain
XX	
XX	Location/Qualifiers
XX	1..21
XX	/note="hydrophobic, transmembrane domain"
XX	
XX	WO9805783-A1.
XX	
XX	12-FEB-1998.
XX	
XX	07-AUG-1997;
XX	97WO-US13945.
XX	
XX	18-MAR-1997;
XX	97US-0040820.
XX	
XX	07-AUG-1996;
XX	96US-0023541.
XX	
XX	18-OCT-1996;
XX	96US-0028515.
XX	
XX	(BIOJ ) BIOGEN INC.
XX	(UYGE-) UNIV GENEVA FACULTY MEDICINE.

PI Browning JL, Chicheportliche Y;  
XX  
XX WPI; 1998-145619/13.  
DR N-PSDB; AAV18599.  
XX  
XX Tumour necrosis factor related ligand - useful for, e.g. treating  
PT cancer, auto-immune disease and immune responses to tissue grafts  
XX  
XX Claim 12; Pages 48-50; 69pp; English.  
PS  
XX The sequence is that of mouse tumour necrosis factor related  
CC ligand (TRELL). TRELL or active fragments can be included with a  
CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
CC diseases or immune responses to tissue grafts, or to stimulate or  
CC suppress the immune system. It is useful to screen for TRELL  
CC receptors, by labelling with a detectable label and screening  
CC compositions for binding. Agents interfering with TRELL-receptor  
CC binding can also be screened for, can then be administered,  
CC optionally with interferon- gamma, to induce cell death or  
CC treat, suppress or alter immune responses (especially involving human  
CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
CC receptor. It's coding sequence can be used in gene therapy for  
CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
CC autoimmune and inflammatory diseases or inherited genetic disorders,  
CC by introducing into cells, and expressing, therapeutically effective  
CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
CC It may also be of use in the preparation of gene encoding probes for  
CC screening natural/synthetic DNAs for TRELL-encoding sequences  
CC and for antisense therapy.  
XX  
XX  
SQ Sequence 225 AA;  
Query Match 100.0%; Score 1162; DB 19; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.3e-112;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VLSIGLALACIGLLVYVSGSNATLSAQPSQBELTAEDRREPPELNPQTEESQDVVPF 60  
DB 1 VLSIGLALACIGLLVYVSGSNATLSAQPSQBELTAEDRREPPELNPQTEESQDVVPF 60  
QY 61 LEQVLRPRRSAPKGRKARPRRAIAAHYEVHPRPGDGAQAGVDGTVSGWETKINSSSPL 120  
DB 61 LEQVLRPRRSAPKGRKARPRRAIAAHYEVHPRPGDGAQAGVDGTVSGWETKINSSSPL 120  
QY 121 RYDRQIGEEFYIRAGLYLYLCQVHFDEGKAVYLLKLDLVNGVLAALRCLLEFSATTAASSPG 180  
DB 121 RYDRQIGEEFYIRAGLYLYLCQVHFDEGKAVYLLKLDLVNGVLAALRCLLEFSATTAASSPG 180  
QY 121 RYDRQIGEEFYIRAGLYLYLCQVHFDEGKAVYLLKLDLVNGVLAALRCLLEFSATTAASSPG 180  
DB 121 RYDRQIGEEFYIRAGLYLYLCQVHFDEGKAVYLLKLDLVNGVLAALRCLLEFSATTAASSPG 180  
QY 181 POLRLCOVSGLLPLRPSSSLRIRLTPWAHLKAAPFLTYFGFQVH 225  
DB 181 POLRLCOVSGLLPLRPSSSLRIRLTPWAHLKAAPFLTYFGFQVH 225  
RESULT 2  
AAB07527  
ID AAB07527 standard; protein; 225 AA.  
XX  
XX AAB07527;  
AC  
XX  
XX 20-OCT-2000 (first entry)  
DT  
XX  
DE Amino acid sequence of a soluble recombinant murine TWEAK protein.  
XX  
XX TWEAK protein; immunological disorder; immune response; inflammation;  
KW TWEAK blocking agent; autoimmune disease; organ transplant rejection;  
KM Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock; tumour.  
XX  
OS Mus sp.  
XX  
XX WO200042073-A1.  
XX  
XX 20-JUL-2000.  
XX  
XX

PF 14-JAN-2000; 2000WO-0501044.  
XX  
XX  
PR 15-JAN-1999; 99US-0116168.  
XX  
XX (BIOI ) BIOGEN INC.  
PA  
XX  
PI Renneft P;  
PI  
XX  
XX WPI; 2000-476036/41.  
DR  
XX  
XX Preventing and treating immune responses using modulators, especially  
PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for  
PT treating e.g. inflammation and graft versus host disease  
XX  
XX Disclosure; Fig 1; 45pp; English.  
PS  
XX The present sequence represents a TWEAK protein. The specification  
CC describes a method for preventing or treating an immunological  
CC disorder and/or inhibiting an immune response in an animal. The  
CC method comprises administering a TWEAK blocking agent. The method may  
CC be used for preventing and treating immune disorders associated with  
CC inappropriate expression and/or activity of TWEAK. These disorders  
CC include autoimmune diseases, acute and chronic inflammation, organ  
CC transplant rejection, Graft-versus-Host disease (GVHD), lymphoid cell  
CC malignancies, septic and other forms of shock, loss of immune  
CC responsiveness (as seen in human immunodeficiency virus (HIV)  
CC infections) and failure of the immune response to tumour growth.  
XX  
XX  
SQ Sequence 225 AA;  
Query Match 100.0%; Score 1162; DB 21; Length 225;  
Best Local Similarity 100.0%; Pred. No. 1.3e-112;  
Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 VLSIGLALACIGLLVYVSGSNATLSAQPSQBELTAEDRREPPELNPQTEESQDVVPF 60  
DB 1 VLSIGLALACIGLLVYVSGSNATLSAQPSQBELTAEDRREPPELNPQTEESQDVVPF 60  
QY 61 LEQVLRPRRSAPKGRKARPRRAIAAHYEVHPRPGDGAQAGVDGTVSGWETKINSSSPL 120  
DB 61 LEQVLRPRRSAPKGRKARPRRAIAAHYEVHPRPGDGAQAGVDGTVSGWETKINSSSPL 120  
QY 121 RYDRQIGEEFYIRAGLYLYLCQVHFDEGKAVYLLKLDLVNGVLAALRCLLEFSATTAASSPG 180  
DB 121 RYDRQIGEEFYIRAGLYLYLCQVHFDEGKAVYLLKLDLVNGVLAALRCLLEFSATTAASSPG 180  
QY 121 RYDRQIGEEFYIRAGLYLYLCQVHFDEGKAVYLLKLDLVNGVLAALRCLLEFSATTAASSPG 180  
DB 121 RYDRQIGEEFYIRAGLYLYLCQVHFDEGKAVYLLKLDLVNGVLAALRCLLEFSATTAASSPG 180  
QY 181 POLRLCOVSGLLPLRPSSSLRIRLTPWAHLKAAPFLTYFGFQVH 225  
DB 181 POLRLCOVSGLLPLRPSSSLRIRLTPWAHLKAAPFLTYFGFQVH 225  
RESULT 3  
AAW93591  
ID AAW93591 standard; protein; 211 AA.  
XX  
XX AAW93591;  
AC  
XX  
XX 18-JUN-1999 (first entry)  
DT  
XX  
XX Mouse TNRL3 protein.  
DE  
XX  
XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;  
KW developmental abnormality; gestational abnormality; prostate cancer;  
KM APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
KM apoptosis; mouse.  
XX  
XX Mus sp.  
XX  
XX WO9911791-A2.  
XX  
XX 11-MAR-1999.  
XX  
XX

PF 04-SEP-1998; 98WO-US18393.  
 XX  
 XX 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNIV ) UNIV WASHINGTON.  
 XX  
 PI Chaudhary PM;  
 XX  
 DR WPI: 1999-205191/17.  
 XX N-PSDB; AAX23425.  
 XX  
 PT New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 XX  
 PS Claim 40; Fig 13B; 156pp; English.  
 XX  
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC moiety. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the changer in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active  
 CC fragments and APO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
 CC activity. The method is performed in vivo or in vitro. APO polypeptides  
 CC are all useful as immunogens for preparing antibodies. APO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. APO8 was transfectected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 XX  
 SQ Sequence 211 AA;  
 Query Match 93.7%; Score 1089; DB 20; Length 211;  
 Best Local Similarity 99.5%; Pred. No. 4.9e-105;  
 Matches 210; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 OY 15 LVVSLGSMATLSAQEPSEELTAEDRREPPLNPOTESQDVPFLDVRPRSPAPG 74  
 DB 1 LVVSLGSMATLSAQEPSEELTAEDRREPPLNPOTESQDVPFLDVRPRSPAPG 60  
 OY 75 RKARRRAIAAHYEVHPRGQDGAQAGVGVSGWEETKINSSPLRYDROIGFEFTVIRA 134  
 DB 61 RKARRRAIAAHYEVHPRGQDGAQAGVGVSGWEETKINSSPLRYDROIGFEFTVIRA 120  
 OY 135 GLIYLYCOVHDEGKAVYIKLIDLNVGVYALRCLEEFSAATASSGPGQLRCQVSGLLPL 194  
 DB 121 GLIYLYCOVHDEGKAVYIKLIDLNVGVYALRCLEEFSAATASSGPGQLRCQVSGLLPL 180  
 OY 195 RFGSSLRIRTLPMALKAAPFLTYGGLFOVH 225  
 DB 181 RFGSSLRIRTLPMALKAAPFLTYGGLFOVH 211  
 RESULT 4  
 AAM29745  
 ID AAM29745 standard; Protein: 249 AA.  
 XX  
 AC AAM29745;  
 XX  
 DT 27-OCT-1998 (first entry)  
 XX  
 DE TNF related endothelium proliferative agent protein.  
 XX  
 KW TNF; endothelium proliferative agent; TREPA; wound healing; cancer;

KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9835061-A2.  
 XX  
 PD 13-AUG-1998.  
 XX  
 PE 12-FEB-1998; 98WO-US02859.  
 XX  
 PR 10-FEB-1998; 98US-0021706.  
 PR 12-FEB-1997; 97US-0798692.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Wiley SR;  
 XX  
 DR WPI: 1998-447255/38.  
 DR N-PSDB; AAV47613.  
 XX  
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 PT treatment of autoimmune disease, tumours and inflammation  
 XX  
 PS Claim 16; Page 123-4; 142pp; English.  
 XX  
 CC The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).  
 XX  
 SQ Sequence 249 AA;  
 Query Match 87.8%; Score 1020; DB 19; Length 249;  
 Best Local Similarity 88.8%; Pred. No. 9.5e-98;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;  
 OY 2 LSLGLALACGLLLVVSIGSWATLSAQEPSEELTAEDRREPPLNPOTESQDVPFL 61  
 DB 26 LGLGLALACGLLLVVSIGSWATLSAQEPSEELTAEDRREPPLNPOTESQDVPFL 85  
 OY 62 RQIVRRSAPRGRRARRRAIAAHYEVHPRGQDGAQAGVGVSGWEETKINSSPLR 121  
 DB 86 RQIVRRSAPRGRRARRRAIAAHYEVHPRGQDGAQAGVGVSGWEETKINSSPLR 145  
 OY 122 YDRQIGFEFTVIRAGIYLYCOVHDEGKAVYIKLIDLNVGVYALRCLEEFSAATASSGPG 181  
 DB 146 YKROIGFEFTVIRAGIYLYCOVHDEGKAVYIKLIDLNVGVYALRCLEEFSAATASSGPG 205  
 OY 182 QLRICQVSGLLPLRPGSSLRIRTLPMALKAAPFLTYGGLFOVH 225  
 DB 206 QLRICQVSGLLPLRPGSSLRIRTLPMALKAAPFLTYGGLFOVH 249  
 RESULT 5  
 AAY09369  
 ID AAY09369 standard; Protein: 249 AA.  
 XX  
 AC AAY09369;  
 XX  
 DT 15-JUL-1999 (first entry)  
 XX  
 DE Human tumour necrosis factor Apo-3 ligand protein sequence.

XX Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;  
 KW NF-kappaB-dependent transcription; JNK/SAPK-dependent response;  
 KM cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09919490-A1.  
 XX  
 PD 22-APR-1999.  
 XX  
 PF 09-OCT-1998; 98WO-US21407.  
 XX  
 PR 17-DEC-1997; 97US-0069862.  
 PR 10-OCT-1997; 97US-0062037.  
 XX  
 PA (GENTECH ) GENENTECH INC.  
 XX  
 PI Ashkenazi AJ, Masters SA, Pitti R;  
 DR WPI: 1999-287982/24.  
 DR N-PSDB: AAX56000.  
 XX  
 PT New human Apo-3 ligand (a tumour necrosis factor) homologue  
 PS  
 PS Claim 1; Fig 1; 74pp; English.  
 XX  
 CC The present sequence represents a human tumour necrosis factor (TNF)  
 CC and lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has  
 CC cytostatic activity. Apo-3 ligand can be used to induce apoptosis in  
 CC mammalian cancer cells, to induce NF-kappaB-dependent transcription and  
 CC to induce JNK/SAPK-dependent responses in mammalian cells.  
 XX  
 SQ Sequence 249 AA;  
 Query Match 87.8%; Score 1020; DB 20; Length 249;  
 Best Local Similarity 88.8%; Pred. No. 9.5e-98;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;  
 QY 2 LSLGLALACGLLLVYVSLGSMATLSAQPSEBELTAEDERPPELNPQTEESODVYVFL 61  
 Db 26 LGGLALACGLLLAVVSLGSRASLSAQEPQELVAEEDDPSELNPQTEESODPAPFL 85  
 QY 62 EQLYRPRRSAPKGRKAPRRARIAAHYEVHPRPGDGAQAGVDGVSGWEETKINSSPLR 121  
 Db 86 NRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGDGAQAGVDGVSGWEETKINSSPLR 145  
 QY 122 YDRQIGFTYIRAGLYLYLCQVHDEGKAVYLKLDLVNGVLAIRCLSEESATASSPGP 181  
 Db 146 YNRQIGFTYIRAGLYLYLCQVHDEGKAVYLKLDLVNGVLAIRCLSEESATASSPGP 205  
 QY 182 QLRICQVSGLLPLRPGSSLRIRTLPMANLKAAPLYTFFGLFQVH 225  
 Db 206 QLRICQVSGLLALRPGSSLRIRTLPMANLKAAPLYTFFGLFQVH 249  
 RESULT 6  
 AAB07526  
 ID AAB07526 standard; protein: 249 AA.  
 XX  
 AC AAB07526;  
 XX  
 DT 20-OCT-2000 (first entry)  
 XX  
 DE Amino acid sequence of a soluble recombinant human TWEAK protein.  
 XX  
 KW TWEAK protein; immunological disorder; immune response; inflammation;  
 KW TWEAK blocking agent; autoimmune disease; organ transplant rejection;  
 KW Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock; tumour.  
 XX  
 OS Homo sapiens.  
 OS  
 PN W0200042073-A1.

XX 20-JUL-2000.  
 PD  
 XX 14-JAN-2000; 2000WO-US01044.  
 PF  
 XX 15-JAN-1999; 99US-0116168.  
 PR  
 XX (BIOJ ) BIOGEN INC.  
 XX  
 PI Renner P;  
 DR WPI: 2000-476036/41.  
 XX  
 PT Preventing and treating immune responses using modulators, especially  
 PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for  
 PT treating e.g. inflammation and graft versus host disease -  
 XX  
 PS Disclosure; Fig 1; 45pp; English.  
 XX  
 CC The present sequence represents a TWEAK protein. The specification  
 CC describes a method for preventing or treating an immunological  
 CC disorder and/or inhibiting an immune response in an animal. The  
 CC method comprises administering a TWEAK blocking agent. The method may  
 CC be used for preventing and treating immune disorders associated with  
 CC inappropriate expression and/or activity of TWEAK. These disorders  
 CC include autoimmune diseases, acute and chronic inflammation, organ  
 CC transplant rejection, Graft-versus-Host disease (GVHD), lymphoid cell  
 CC malignancies, septic and other forms of shock, loss of immune  
 CC responsiveness (as seen in human immunodeficiency virus (HIV)  
 CC infections) and failure of the immune response to tumour growth.  
 XX  
 SQ Sequence 249 AA;  
 Query Match 87.8%; Score 1020; DB 21; Length 249;  
 Best Local Similarity 88.8%; Pred. No. 9.5e-98;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;  
 QY 2 LSLGLALACGLLLVYVSLGSMATLSAQPSEBELTAEDERPPELNPQTEESODVYVFL 61  
 Db 26 LGGLALACGLLLAVVSLGSRASLSAQEPQELVAEEDDPSELNPQTEESODPAPFL 85  
 QY 62 EQLYRPRRSAPKGRKAPRRARIAAHYEVHPRPGDGAQAGVDGVSGWEETKINSSPLR 121  
 Db 86 NRLVRPRRSAPKGRKTRARRAIAAHYEVHPRPGDGAQAGVDGVSGWEETKINSSPLR 145  
 QY 122 YDRQIGFTYIRAGLYLYLCQVHDEGKAVYLKLDLVNGVLAIRCLSEESATASSPGP 181  
 Db 146 YNRQIGFTYIRAGLYLYLCQVHDEGKAVYLKLDLVNGVLAIRCLSEESATASSPGP 205  
 QY 182 QLRICQVSGLLPLRPGSSLRIRTLPMANLKAAPLYTFFGLFQVH 225  
 Db 206 QLRICQVSGLLALRPGSSLRIRTLPMANLKAAPLYTFFGLFQVH 249  
 RESULT 7  
 AAY95338  
 ID AAY95338 standard; Protein: 249 AA.  
 XX  
 AC AAY95338;  
 XX  
 DT 25-SEP-2000 (first entry)  
 XX  
 DE Human PRO207 antitumour protein.  
 XX  
 KW PRO207; human; antitumour; tumour; therapy; cytostatic;  
 KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
 KW uterine cancer; prostate cancer; lung cancer; bladder cancer;  
 KW central nervous system cancer; melanoma; leukaemia; neoplasm.  
 XX  
 OS Homo sapiens.  
 OS  
 PN Key Location/Qualifiers  
 FH Peptide 1..40  
 FT

FT Protein /label=Signal\_peptide  
 FT 41..249 /label=PRO207  
 FT Modified-site 27..33 /note="N-myristoylation"  
 FT Modified-site 29..35 /note="N-myristoylation"  
 FT Modified-site 36..42 /note="N-myristoylation"  
 FT Modified-site 45..51 /note="N-myristoylation"  
 FT Modified-site 118..124 /note="N-myristoylation"  
 FT Modified-site 121..127 /note="N-myristoylation"  
 FT Modified-site 125..131 /note="N-myristoylation"  
 FT Modified-site 128..134 /note="N-myristoylation"  
 FT Modified-site 139..143 /note="N-myristoylation"  
 FT Modified-site 10..14 /note="Asn is N-glycosylated"  
 FT Modified-site 97..101 /note="amidation"  
 FT Modified-site 24..35 /note="amidation"  
 FT Peptide /note="prokaryotic membrane lipoprotein lipid"  
 FT WO200037638-A2.  
 XX 29-JUN-2000.  
 PD 02-DEC-1999; 99WO-US28565.  
 XX 22-DEC-1998; 98US-013296.  
 PR 08-MAR-1999; 99WO-US05028.  
 PR 21-APR-1999; 99US-0130232.  
 PR 28-APR-1999; 99US-0131445.  
 PR 14-MAY-1999; 99US-0134287.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 15-SEP-1999; 99WO-US21090.  
 PR 15-SEP-1999; 99WO-US21547.  
 XX (GENH ) GENENTECH INC.  
 PA Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;  
 PI Napier MA, Pilti RM, Wood WI;  
 XX N-PSDB; AAA49717.  
 DR WPI: 2000-442668/38.  
 XX Novel composition to inhibit neoplastic cell growth or for treating  
 PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,  
 PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or  
 PT PRO866  
 XX Claim 19; Fig 4; 172pp; English.  
 PS  
 XX The present sequence is that of human antitumour protein PRO207,  
 CC as deduced from a foetal kidney cDNA clone (see AAA49717). PRO207  
 CC shows amino acid sequence identity to tumour necrosis factor family  
 CC members, especially human lymphotoxin-beta (23.4%) and human CD40  
 CC ligand (19.8%). Mol.wt. is 27,216. A claimed method for inhibiting  
 CC the growth of a tumour cell comprises exposing the tumour cell  
 CC to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224, PRO328, PRO301,  
 CC PRO526, PRO362, PRO509 or PRO866 (see AAY95337-49), their  
 CC agonists or chimeric polypeptides incorporating them. The tumour  
 CC is especially a cancer selected from breast, ovarian, renal,  
 CC colorectal, uterine, prostate, lung, bladder and central nervous  
 CC system cancer, melanoma and leukaemia. Methods for the recombinant  
 CC expression of the antitumour proteins are also provided.

SQL Sequence 249 AA;  
 Query Match 87.8%; Score 1020; DB 21; Length 249;  
 Best Local Similarity 88.8%; Pred. No. 9.5e-98;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;  
 QY 2 LSLGLALACGLLVVSWATLSAOPSEETADREPPPELNQTESQDVPFL 61  
 DB 26 LGLGLALACGLLVVSWATLSAOPSEETADREPPPELNQTESQDVPFL 85  
 QY 62 EQLVPRRSAPKGRARRRAIAHYEVHPRGQGAQVGVTSWGEETKINSSEPLR 121  
 DB 86 NKLVPRRSAPKGRARRRAIAHYEVHPRGQGAQVGVTSWGEETKINSSEPLR 145  
 QY 122 YDRQGEFTVIRAGLYLYICQVHPEGRKAVYIKLDLVGVATLCLFEFSSTAASSRCP 181  
 DB 146 YNRQGEFTVIRAGLYLYICQVHPEGRKAVYIKLDLVGVATLCLFEFSSTAASSRCP 205  
 QY 182 QLRLQVSGILPLRPGSSLRIRTLPMHLKAPFLTYFGLFQVH 225  
 DB 206 QLRLQVSGILALRRGSSLRIRTLPMHLKAPFLTYFGLFQVH 249  
 RESULT 8  
 ID AAE00891 standard; Protein: 249 AA.  
 AC AAE00891;  
 DE 04-JUL-2001 (first entry)  
 DE Human TREPA (TNF related endothelium proliferative agent).  
 KW Human: tumour necrosis factor; TNF; angiogenesis; wound healing;  
 KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;  
 KW grafting; vuneratary.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Domain 98..249 /label= Extracellular\_domain  
 FT US6207642-B1.  
 PN 27-MAR-2001.  
 PD 26-JUN-1998; 98US-0105343.  
 PR 12-FEB-1997; 97US-0798692.  
 PR 10-FEB-1998; 98US-0021706.  
 XX (ABBO ) ABBOTT LAB.  
 PA Wiley SR;  
 PI WPI: 2001-280760/29.  
 DR N-PSDB; AAD04350.  
 XX Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 PT -  
 XX Claim 1; Column 75-76; 53pp; English.  
 PS  
 XX The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts.







KM ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 KM retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 KM rubosis; uveitis; macular degeneration; arthritis; rheumatism;  
 KM corneal graft neovascularisation; psoriasis; metastatic condition;  
 KM malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
 KM preneoplastic condition; myocardial angiogenesis; wound granulation;  
 KM scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
 KM atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 KM peripheral atherosclerosis; pDC409-L2-TWEAK; TWEAK receptor; TWEAKR;  
 KM fusion protein.

OS Homo sapiens.  
 OS Synthetic.

PN W0200145730-A2.

PD 28-JUN-2001.

PF 19-DEC-2000; 2000WO-US34755.

PR 20-DEC-1999; 98US-0172878.

PR 10-MAY-2000; 2000US-0203347.

PA (IMV ) IMMUNEX CORP.

PI Wiley SR;

DR WPI: 2001-417975/44.

DR N-PSDB: AAS03964.

PT Modulating angiogenesis in a mammal for treating diseases mediated by  
 PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  
 PT peripheral tissue, by administering antagonist or agonist of TWEAK  
 PT receptor

PS Example 1; Page 41; 46pp; English.

CC The sequence represents a fusion protein encoded by the  
 CC expression vector pDC409-L2-TWEAK. The fusion protein comprises a  
 CC growth hormone leader, a leucine zipper multimerisation domain, and  
 CC the extracellular domain of human TWEAK. The fusion protein was  
 CC used in the isolation of human TWEAK receptor (TWEAKR)-expressing  
 CC clones from a COS cell human cDNA library. The TWEAK protein is  
 CC a member of the tumour necrosis factor (TNF) family and induces  
 CC angiogenesis. TWEAKR may therefore be used to screen for and  
 CC develop TWEAKR agonists and antagonists for the modulation of  
 CC angiogenesis, to be used in the treatment and diagnosis of human disease.  
 CC The disorders mediated by angiogenesis include ocular disorders  
 CC characterised by ocular neovascularisation such as diabetic retinopathy,  
 CC neovascular glaucoma, retinoblastoma, retinopathy of prematurity,  
 CC retrolental fibroplasia, rubosis, uveitis, macular degeneration and  
 CC corneal graft neovascularisation, and inflammatory diseases such as  
 CC arthritis, rheumatism and psoriasis. Other treatable diseases include  
 CC malignant and metastatic conditions such as sarcomas and carcinomas,  
 CC benign tumours and preneoplastic conditions, myocardial angiogenesis,  
 CC haemophilic joints, scleroderma, vascular adhesions, atherosclerotic  
 CC plaque neovascularisation, telangiectasia, wound granulation, coronary  
 CC atherosclerosis, peripheral atherosclerosis and ischaemia.

SQ Sequence 273 AA;

Query Match 81.8%; Score 951; DB 22; Length 273;

Best Local Similarity 88.9%; Pred. No. 1.6e-90;

Matches 184; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

OY 19 SLGSMATLSAQEPQOEELTADEPPELNPQTESQDVVFLQOLVPRRSAPKGRAR 78

DB 67 SLGSRASLSAQEPQOEELVAEDDPSELNQTESQDPAPFLNLRPRRSAPKGRKTR 126

OY 79 PRRAIAHYEYHPRPGDGAQAGVDGTVSGWEETKINSSEPLRYDRQIGFTYRAGLY 138

DB 127 ARRAIAHYEYHPRPGDGAQAGVDGTVSGWEETKINSSEPLRYDRQIGFTYRAGLY 186

OY 139 LYCQVHFDEGKAVYIKLIDLNVGVALRCLIEFSATMASSPGPOLRLCOVSGILPLRPS 198  
 DB 187 LYCQVHFDEGKAVYIKLIDLNVGVALRCLIEFSATMASSPGPOLRLCOVSGILPLRPS 246

OY 199 SIRIRTLPAHKLKAAPELTYFGLFOVH 225

DB 247 SLIRIRTLPAHKLKAAPELTYFGLFOVH 273

RESULT 13  
 AAM29746  
 ID AAM29746 standard; Protein; 189 AA.

AC AAM29746;

DE 27-OCT-1998 (first entry)

DE TNF related endothelium proliferative agent protein 2.

KM TNF; endothelium proliferative agent; TREPA; wound healing; cancer;  
 KM tissue grafting; vascularisation; apoptosis; autoimmune; birth control.

OS Homo sapiens.

PN W09835061-A2.

PD 13-AUG-1998.

PF 12-FEB-1998; 98WO-US02859.

PR 10-FEB-1998; 98US-0021706.

PR 12-FEB-1997; 97US-0798692.

PA (ABBO ) ABBOTT LAB.

PI Wiley SR;

DR WPI: 1998-447255/38.

PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 PT treatment of autoimmune disease, tumours and inflammation

PS Claim 16; Page 125-6; 142pp; English.

CC The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).

SQ Sequence 189 AA;

Query Match 63.9%; Score 742; DB 19; Length 189;

Best Local Similarity 92.8%; Pred. No. 5.7e-69;

Matches 142; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 73 KGRARPRRAIAHYEYHPRPGDGAQAGVDGTVSGWEETKINSSEPLRYDRQIGFTY 132

DB 37 KGRTRARRRGIAIAHYEYHPRPGDGAQAGVDGTVSGWEETKINSSEPLRYDRQIGFTY 96

OY 133 RAGLYLYYLCQVHFDEGKAVYIKLIDLNVGVALRCLIEFSATMASSPGPOLRLCOVSGIL 192

DB 97 RAGLYLYYLCQVHFDEGKAVYIKLIDLNVGVALRCLIEFSATMASSPGPOLRLCOVSGIL 156

```
OY 193 PLRPGSSLRIRLTPMAHLKAAPFLTYFGLFOVH 225
      |||||||
DB 157 ALRPGSSLRIRLTPMAHLKAAPFLTYFGLFOVH 189

RESULT 14
AAE00892
ID AAE00892 standard; Protein; 189 AA.
XX
AC AAE00892;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human ULAflag TREPA soluble construct.
XX
KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing; tumour;
KW TREPA; TNF related endothelium proliferative agent; metastasis; grafting;
KW vulnereary; HUVEC; human umbilical vein endothelial cell; ULAflag.
XX
OS Homo sapiens.
XX
PN US6207642-B1.
XX
PD 27-MAR-2001.
XX
PF 26-JUN-1998; 98US-0105343.
XX
PR 12-FEB-1997; 97US-0798692.
PR 10-FEB-1998; 98US-0021706.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Wiley SR;
XX
DR WPI; 2001-280760/29.
XX
PT Inducing angiogenesis in mammal at desired sites for promoting wound
PT healing, by administering soluble fragment of extracellular domain of
PT tumor necrosis factor related endothelium proliferative agent protein
PT
XX
XX Example 2; Column 75-78; 53pp; English.
XX
CC The present invention relates to extracellular signal molecules,
CC particularly members of tumour necrosis factor (TNF) family molecules
CC designated as TREPA (TNF related endothelium proliferative agent).
CC Soluble biologically active TREPA are used to treat TREPA-associated
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis
CC in human for promoting wound healing and for vascularising grafted
CC tissue for successful grafting and to promote tissue grafts. The present
CC amino acid sequence is human ULAflag TREPA soluble construct. This
CC sequence which is a biologically active molecule is capable of inducing
CC proliferation in HUVEC (human umbilical vein endothelial cells) cells.
XX
SQ Sequence 189 AA;

Query Match 63.9%; Score 742; DB 22; Length 189;
Best Local Similarity 92.8%; Pred. No. 5.7e-69;
Matches 142; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

OY 73 KGRKARPRRAIAHYEVHPRPGODGAQAGVDGTVSGMEETKINSSPLRYDRQIGFTYI 132
      |||||
DB 37 KGRKTRARGIAHYEVHPRPGODGAQAGVDGTVSGMEARINSSPLRYNQGIFITYT 96

OY 133 RAGLYLYTCQVHFDEGKAVYLLKLDLVNGVIALRCLIEFSATPAASPGQRLICQVSGLL 192
      |||||||
DB 97 RAGLYLYTCQVHFDEGKAVYLLKLDLVNGVIALRCLIEFSATPAASISGQRLICQVSGLL 156

OY 193 PLRPGSSLRIRLTPMAHLKAAPFLTYFGLFOVH 225
      |||||||
DB 157 ALRPGSSLRIRLTPMAHLKAAPFLTYFGLFOVH 189
```

```
RESULT 15
AAE00895
ID AAE00895 standard; Protein; 146 AA.
XX
AC AAE00895;
XX
DT 04-JUL-2001 (first entry)
XX
DE Human TREPA (TNF related endothelium proliferative agent) fragment.
XX
KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing;
KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;
KW grafting; vulnereary.
XX
OS Homo sapiens.
XX
PN US6207642-B1.
XX
PD 27-MAR-2001.
XX
PF 26-JUN-1998; 98US-0105343.
XX
PR 12-FEB-1997; 97US-0798692.
PR 10-FEB-1998; 98US-0021706.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Wiley SR;
XX
DR WPI; 2001-280760/29.
XX
PT Inducing angiogenesis in mammal at desired sites for promoting wound
PT healing, by administering soluble fragment of extracellular domain of
PT tumor necrosis factor related endothelium proliferative agent protein
PT
XX
XX Example 14; Fig.1; 53pp; English.
XX
CC The present invention relates to extracellular signal molecules,
CC particularly members of tumour necrosis factor (TNF) family molecules
CC designated as TREPA (TNF related endothelium proliferative agent).
CC Soluble biologically active TREPA are used to treat TREPA-associated
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis
CC in human for promoting wound healing and for vascularising grafted
CC tissue for successful grafting and to promote tissue grafts.
CC The present amino acid sequence is human TREPA fragment.
XX
SQ Sequence 146 AA;

Query Match 62.0%; Score 721; DB 22; Length 146;
Best Local Similarity 94.5%; Pred. No. 6.1e-67;
Matches 138; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 80 RRAIAHYEVHPRPGODGAQAGVDGTVSGMEETKINSSPLRYDRQIGFTYIRAGLYYL 139
      |||||||
DB 1 RRAIAHYEVHPRPGODGAQAGVDGTVSGMEARINSSPLRYNQGIFITYRAGLYYL 60

OY 140 YCQVHFDEGKAVYLLKLDLVNGVIALRCLIEFSATPAASPGQRLICQVSGLLPLRPGSS 199
      |||||||
DB 61 YCQVHFDEGKAVYLLKLDLVNGVIALRCLIEFSATPAASISGQRLICQVSGLLALRPGSS 120

OY 200 LRIRLTPMAHLKAAPFLTYFGLFOVH 225
      |||||||
DB 121 LRIRLTPMAHLKAAPFLTYFGLFOVH 146

Search completed: May 8, 2003, 02:19:41
Job time : 29.7328 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 07:13:39 ; Search time 1561.93 Seconds  
(without alignments)  
12110.869 Million cell updates/sec

Title: US-09-245-198a-1  
Perfect score: 1168  
Sequence: 1 ggtgtgcgacctgagcctg.....ataaatcatgattctcttc 1168

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_estba:\*  
3: em\_estba:\*  
4: em\_estba:\*  
5: em\_estba:\*  
6: em\_estba:\*  
7: em\_estba:\*  
8: em\_estba:\*  
9: em\_estba:\*  
10: em\_estba:\*  
11: em\_estba:\*  
12: em\_estba:\*  
13: em\_estba:\*  
14: em\_estba:\*  
15: em\_estba:\*  
16: em\_estba:\*  
17: em\_estba:\*  
18: em\_estba:\*  
19: em\_estba:\*  
20: em\_estba:\*  
21: em\_estba:\*  
22: em\_estba:\*  
23: em\_estba:\*  
24: em\_estba:\*  
25: em\_estba:\*  
26: em\_estba:\*  
27: em\_estba:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	883.8	75.7	1033	11	AK020909 Mus musculus
2	625.8	53.6	918	12	BF577781 BF577781
3	519.4	44.5	731	13	BI871711 BI871711
4	507.4	43.4	728	13	BI870393 BI870393
5	504	43.2	561	10	AM763237 AM763237
6	497	42.6	533	10	BE628951 BE628951

7	489.6	41.9	650	12	BG404836	BG404836 602240160
8	488.8	41.8	587	14	BO208433	BO208433 UT-R-EP0
9	481.2	41.2	584	10	AM917574	AM917574 EST348878
10	480.4	41.1	777	13	BI819200	BI819200 603034614
11	473	40.5	834	13	BI766766	BI766766 603056866
12	467.8	40.1	948	14	BO707185	BO707185 AGENCOURT
13	456.4	39.1	828	13	BI596681	BI596681 603243254
14	437.8	37.5	471	9	AA221610	AA221610 my18d09.r
15	421.6	36.1	963	14	BO671259	BO671259 AGENCOURT
16	419.2	35.9	940	14	BO884231	BO884231 AGENCOURT
17	406.4	34.8	1071	14	BM921213	BM921213 AGENCOURT
18	390.6	33.4	445	9	AA870722	AA870722 vq25907.r
19	366.6	31.4	951	14	BO674188	BO674188 AGENCOURT
20	360	30.8	360	10	BE654876	BE654876 UT-M-BHO-
21	353	30.2	405	9	AT854476	AT854476 UT-M-BHO-
22	331	28.3	367	13	BI111534	BI111534 602895883
23	325.4	27.9	376	9	AA792068	AA792068 vns6d09.r
24	309	26.5	474	13	BI965174	BI965174 1d34e07.y
25	306	26.2	315	12	BE466521	BE466521 UT-M-CGDP
26	299.4	25.6	538	12	BF821434	BF821434 MRI-RT003
27	297.2	25.4	400	9	AT152313	AT152313 ud87h02.r
28	292.8	25.1	493	10	BE307031	BE307031 601087888
29	289.6	24.8	433	12	BF283688	BF283688 EST448279
30	289.4	24.8	542	12	BF041509	BF041509 BP2500258
31	285.2	24.4	443	12	BG378802	BG378802 UT-R-CV1-
32	284.4	24.3	785	13	BI762908	BI762908 603047966
33	283.8	24.3	404	12	BF044430	BF044430 BP250013A
34	278.4	23.8	897	13	BI730298	BI730298 603350276
35	274.4	23.5	412	9	AT101416	AT101416 EST204867
36	271	23.2	581	13	BI738634	BI738634 603358846
37	270.4	23.2	402	12	BF410871	BF410871 UT-R-CNO-
38	270.2	23.1	292	10	BB268794	BB268794 BB268794
39	258.2	22.1	558	13	BM494863	BM494863 538865 MA
40	255.2	21.8	394	12	EG376757	EG376757 UT-R-CU0-
41	246.4	21.1	298	14	BE688946	BE688946 UT-E-CQ1-
42	243.6	20.9	371	13	BI300370	BI300370 UT-R-CV2-
43	242.4	20.8	441	13	BI967060	BI967060 1d27h07.y
44	242.2	20.7	374	9	AA800970	AA800970 EST190467
45	240	20.5	240	10	AM764050	AM764050 ut70d09.x

## ALIGNMENTS

RESULT 1	AK020909	1033 bp	mRNA	linear	HTC 19-JAN-2002
LOCUS	AK020909				
DEFINITION	Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone:A930030D13:tumor necrosis factor (ligand) superfamily, member 12, full insert sequence.				
ACCESSION	AK020909	1	GI:12861640		
VERSION	AK020909.1				
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (strain:C57BL/6J) adult retina cDNA to mRNA, clone:lib:RIKEN full-length enriched mouse cDNA library clone:A930030D13.				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
POBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20493374				
PUBMED	11042159				



RESULT 2	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Db	422	CANSTACTAG66GCGCTTGTCTCTCCACAGATTCCTTAACTTTTCCCGGCTCCAGAGAC	481									
Qy	728	ATCACACACACCTCCCTACCCACCCACCCTCCTCCACCCCTCGCTGCTCTTGGTCCAG	787									
Db	482	ATCACACACACCTTTCATCCACCACCCACCTCCTCCACCCCTAGCTGCTTGGTCCAG	541									
Qy	788	TCCGCT-CTCCGCTCAAGGACGACGAGCTTTGTCATATG-TTTCATTTCCACGACGT	845									
Db	542	TCCGCTCTCTCTCTCAAGGACGACGAGCTTTGTCATATGTTTCCATTTCCACGACGT	601									
Qy	846	ATCCCTGCTCTTC-TTAACTATCCATCCACACACAACTATCCACTCTACTAGTCCCAA	904									
Db	602	ATCCCTGCTCTCTTTAACTATCCATCCACACACAACTATCCACTCTACTAGTCCCAA	661									
Qy	905	GCCCCCTACTTATCCCTGACTCCCCACCCTCCTACCCGACACAGCTGTTTATGACTTTGT	964									
Db	662	GCCCCCTACTTATCCCTGACTCCCCACCCTCCTACCCGACACAGCTGTTTATGACTTTGT	721									
Qy	965	GCACACGACCTGAGATGGGTGGACCTGGTGGACGAGAAAGCAGAACTGGGACTAGG	1024									
Db	722	GCACACGACCTGAGATGGGTGGACCTGGTGGACGAGAAAGCAGAACTGGGACTAGG	781									
Qy	1025	CCAGAACTTCCCACTGTGAGGGGGGAAAGCTGGGAGAAAGCTCCTCCCTGGATCCCTGT	1084									
Db	782	CCAGAACTTCCCACTGTGAGGGGGGAAAGCTGGGAGAAAGCTCCTCCCTGGATCCCTGT	841									
Qy	1085	GGATTTTGAAGAAAGATCTATTTTATTTATTTATTTATTTGACAAATGTTAAATGGATTTAA	1144									
Db	842	GGATTTTGAAGAAAGATCTATTTTATTTATTTATTTATTTGACAAATGTTAAATGGATTTAA	901									
Qy	1145	GAGATTAATCATGATTTCTCTC 1158										
Db	902	GAGATTAATCATGATTTCTCTC 925										
RESULT 2	BF577781	918 bp mRNA linear EST 12-DEC-2000										
LOCUS	BF577781	602092080F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4206595 5'										
DEFINITION	BF577781	mRNA sequence.										
ACCESSION	BF577781											
VERSION	BF577781.1	GI:11651493										
KEYWORDS		EST.										
SOURCE		house mouse.										
ORGANISM		Mus musculus										
REFERENCE		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;										
AUTHORS		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.										
TITLE		1 (bases 1 to 918)										
JOURNAL		NIH-MGC htlp://mgc.nci.nih.gov/.										
COMMENT		National Institutes of Health, Mammalian Gene Collection (MGC)										
		Unpublished (1999)										
		Contact: Robert Strausberg, Ph.D.										
		Email: cgapbs-remail.nih.gov										
		Tissue Procurement: Jeffrey E. Green, M.D.										
		cDNA Library Preparation: Life Technologies, Inc.										
		CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)										
		DNA sequencing by: Incyte Genomics, Inc.										
		Clone distribution: MGC clone distribution information can be										
		found through the I.M.A.G.E. Consortium/LNLN at:										
		http://image.lnl.gov										
		Plate: LLM9767 row: p column: 20										

		Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCICGP Library."
BASE COUNT	153 a 292 c 268 g 204 t	1 others
ORIGIN		
Query Match	53.6%; Score 625.8; DB 12; Length 918;	
Best Local Similarity	99.2%; Pred. No. 7.2e-152;	
Matches 660; Conservative 0; Mismatches 2; Indels 3; Gaps 3;		
OY	22 GCTGGCGTCCCTTGAGCCTCTGCTGTGTGTGTACAGCTGGGGAGCTGGGCACAACGCTGTC	81
Dd	1 GGTGGCTGCGCTTGCGCTCTGCTGTGTGTGTACGCTGGGGAGCTGGGCACAACGCTGTC	60
OY	82 TGCCAGAGAGCCTTTCAAGAGAGAGCTGACACAGAAGACC GCCGGAGGCCCTCTGAATCT	141
Dd	61 TGCCAGAGAGCCTTTCAAGAGAGAGCTGACACAGAAGACC GCCGGAGGCCCTCTGAATCT	120
OY	142 GAATCCCCAGACAGAGAAAGCCAGAGTAGTGTTACTCTTCTTGAACAATACTATGTCGGGC	201
Dd	121 GAATCCCCAGACAGAGAAAGCCAGAGTAGTGTTACTCTTCTTGAACAATACTATGTCGGGC	180
OY	202 TCGAAGAAGTGTCTCTTAAGGCCCGGAAGGCCGCCCTGCGCAGCTATTGACCCCATTA	261
Dd	181 TCGAAGAAGTGTCTCTTAAGGCCCGGAAGGCCGCC - CTGGCCAGGTATTGACGCCCATTA	239
OY	262 TGAGTTATCTCTGGCCAGACAGAGATGAGCACACAGCAGSTGTGATGGACAGTGAG	321
Dd	240 TGAGTTATCTCTCTGGCCAGACAGAGATGAGCACACAGCAGSTGTGATGGAGACAGTGAG	299
OY	322 TGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCTGCGCTACAGCCGACATATGG	381
Dd	300 TGGCTGGGAAGAGACCAAAATCAACAGCTCCAGCCCTCTGCGCTACAGCCGACATATGG	359
OY	382 GGAATTTACAGTACAGAGGCTGGGCTCTACTACCTGTACTGTACAGTGCACATTTGATGA	441
Dd	360 GGAATTTACAGTACAGAGGCTGGGCTCTACTACCTGTACTGTACAGTGCACATTTGATGA	419
OY	442 GGGAAAAGCTGTCTACTGTGAAGCTGTGACTGTGTGAAAGGTGTGCTGGCCCTGCGCTG	501
Dd	420 GGGAAAAGCTGTCTACTGTGAAGCTGTGACTGTGTGAAAGGTGTGCTGGCCCTGCGCTG	479
OY	502 CCTGGAATAATTCCTAGACACACAGCAAAGCTCTCTGGGCCCAAGCTCCCTTTGTGCA	561
Dd	480 CCTGGAATAATTCCTAGACACACAGCAAAGCTCTCTGGGCCCAAGCTCCCTTTGTGCA	539
OY	562 GGTTGCTGGGCTGTGGCGCTCGGACAGAGGCTTCCCTTGGATCCGACCCCTCCCTG	621
Dd	540 GGTTGCTGGGCTGTGGCGCTCGGACCA - GGCTTTCCCTTCGGAATCCGACCCCTCCCTG	598
OY	622 GGCATCATCTTAAGGCTGCCCTTCTCTAACCTACTTGTGACCTTTCAAGTTCACTGAGG	681
Dd	599 GGCATCATCTTAAGGCTG - CCCTTCTCTAACCTTGTGACCTTTCAAGTTCACTGAGG	657
OY	682 GGCTT 686	
Dd	658 GGCTT 662	
RESULT 3		
B1871711	B1871711	731 bp mRNA linear EST 11-OCT-2001
LOCUS	603395825P1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405478 5'	
DEFINITION	mRNA sequence.	
ACCESSION	B1871711	
VERSION	B1871711.1 GI:16045386	
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 731) NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC)	
TITLE		





Db 241 AGCCATTATGAGTTCATCCAGACGCTGACAGGAGCGGACGACGAGTGTGACGG 300  
 QY 313 GACAGTGGTGGCTGGAGAGAGCCAAATCAACAGCTCCAGCCCTGCGCTAGACGG 372  
 Db 301 GACAGTGGTGGCTGGAGAGAGCCAAATCAACAGCTCCAGCCCTGCGCTAGACGG 360  
 QY 373 CCAGATTGGGGAATTACAGTATCATCAGGCTGGGCTCTACTACTGCTAGGTGCA 432  
 Db 361 CCAGATTGGGGAATTATATGATCACCAGGCTGGGCTCTACTACTGCTAGGTGCA 420  
 QY 433 CTTTGTATGAGGAAAGGCTGTCTACCTGAGCTGAGCTGCTGCTGCTGCTGCTG 492  
 Db 421 CTTTGTATGAGGAAAGGCTGTCTACCTGAGCTGAGCTGCTGCTGCTGCTGCTG 480  
 QY 493 CCGGCTGCTGAGAAATTCACAGCCACAGCAAGCTCTCCGCGGCGCCAGCTCG 552  
 Db 481 CCGGCTGCTGAGAAATTCACAGCCACAGCAAGCTCTCCGCGGCGCCAGCTCG 540  
 QY 553 TTTGTGCGAGTGTCTGGGCTGTGGCGGCGGCGGAGGCTCTCCCTTCGAGTCCGAC 612  
 Db 541 CCGTGGCCAGGTGTCTGGGCTGTGGCGGCGGCGGAGGCTCTCCCTTCGAGTCCGAC 599  
 QY 613 CCGTGGCCAGGTGTCTGGGCTGTGGCGGCGGCGGAGGCTCTCCCTTCGAGTCCGAC 672  
 Db 600 CCGTGGCCAGGTGTCTGGGCTGTGGCGGCGGCGGAGGCTCTCCCTTCGAGTCCGAC 658  
 QY 673 TCAGTGGGCGGCTGCTCCGAG 697  
 Db 659 TCAGTGGGCGGCTGCTCCGAG 683

RESULT 5  
 AW763237 561 bp mRNA linear EST 04-MAY-2000  
 LOCUS ur70d09.y1 NCI-CGAP\_Mam3 Mus musculus cDNA clone IMAGE:3155633 5'  
 DEFINITION similar to TR:054907 054907 TNF-RELATED WEAK INDUCER OF APOPTOSIS  
 ; mRNA sequence.

ACCESSION AW763237  
 VERSION AW763237.1 GI:7695174  
 KEYWORDS EST.  
 SOURCE house mouse.

ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 561)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
 Tumor Gene Index

JOURNAL Unpublished (1997)  
 COMMENT Other\_ESTs: ur70d09.x1  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov

Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 image.llnl.gov/image/html/lresources.shtml

FEATURES  
 SOURCE MGI:1058389  
 Seq primer: -40RP from glbco  
 High quality sequence start: 433.  
 Location/Qualifiers  
 1..561  
 /organism="Mus musculus"  
 /strain="129/C57BL/6J,FVB/N"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:3155633"  
 /clone\_id="NCI-CGAP\_Mam3"  
 /tissue\_type="tumor, gross tissue"  
 /dev\_stage="10 months"  
 /lab\_host="DH10B"

/note="Organ: mammary; Vector: pCMV-Sport6; Site.1: SalI;  
 Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Library constructed by Life Technologies. Investigators  
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH  
 Reference for transgenic model: Xu et al., Nature Genetics  
 22, 37-43 (1999)."  
 BASE COUNT 108 a 158 c 194 g 100 t 1 others  
 ORIGIN

Query Match 43.2%; Score 504; DB 10; Length 561;  
 Best Local Similarity 99.8%; Pred. No. 3.1e-120;  
 Matches 504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGTGTGAGCCCTGGGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 60  
 Db 57 GGTGTGAGCCCTGGGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 116  
 QY 61 GGGGAGCTGGGCGAAGCTGTCTGCCAGAGGCTTCTCAGAGAGCTGACAGCAGAGA 120  
 Db 117 GGGGAGCTGGGCGAAGCTGTCTGCCAGAGGCTTCTCAGAGAGCTGACAGCAGAGA 176  
 QY 121 CCGCGGGAGGCCCTGAATCAGATCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
 Db 177 CCGCGGGAGGCCCTGAATCAGATCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAG 236  
 QY 181 CTTGGAACAATAGTCCGCGCTGGAAGAAGTCTCTAAAGGCGGGAAGGCGGCGCTCG 240  
 Db 237 CTTGGAACAATAGTCCGCGCTGGAAGAAGTCTCTAAAGGCGGGAAGGCGGCGCTCG 296  
 QY 241 CCGAGCTATTGACGCCATTATGAGGTTTCATCTCGGCGGCGGAGAGAGAGAGAGAG 300  
 Db 297 CCGAGCTATTGACGCCATTATGAGGTTTCATCTCGGCGGCGGAGAGAGAGAGAGAG 356  
 QY 301 AGGTGTGATGGGAG 360  
 Db 357 AGGTGTGATGGGAG 416  
 QY 361 GCGCTACGACGCCAGATGGGGAATTTACATCATCAGAGGCTGGGCTCTACTACTGTA 420  
 Db 417 GCGCTACGACGCCAGATGGGGAATTTACATCATCAGAGGCTGGGCTCTACTACTGTA 476  
 QY 421 CTGTGAGGTGACCTTTGATGAGGAAAGGCTGTCTACTGAAAGCTGAGCTGCTGTA 480  
 Db 477 CTGTGAGGTGACCTTTGATGAGGAAAGGCTGTCTACTGAAAGCTGAGCTGCTGTA 536  
 QY 481 CCGTGTGCTGGGCGCTGCGCTGCTG 505  
 Db 537 CCGTGTGCTGGGCGCTGCGCTGCTG 561

RESULT 6  
 BE628951 533 bp mRNA linear EST 25-AUG-2000  
 LOCUS uc30c03.y1 Soares\_mammary\_gland.MMLMG Mus musculus cDNA clone  
 DEFINITION IMAGE:3373444 5' similar to TR:054907 054907 TNF-RELATED WEAK  
 INDUCER OF APOPTOSIS ; mRNA sequence.

ACCESSION BE628951  
 VERSION BE628951.1 GI:9911639  
 KEYWORDS EST.  
 SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 533)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 This clone is available royalty-free through LLNL; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 MGI:1083048



Db	564	ATCCAF-CCNAAGACTATACCTAGATCTACTTAAATAGCCATGCAACAAATATACACA	619
RESULT 8			
BO208433/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
COMMENT			
FEATURES			
SOURCE			
BASE COUNT			
ORIGIN			
Query Match	41.88;	Score 488.8;	DB 14; Length 687;
Best Local Similarity	87.96;	Pred. No. 3.1e-116;	

[illegible]



```

Db 434 TCCCTGGGGCCCAAGCTCCGCTGCGAGGTGTGGGCTTGGGCGCCCTGGGCGAGG 493
QY 593 TCTTCCCTTGGATCCGACACCTCCCTGGGCTATCTTAAGCTGCCCTTCCATAC 652
Db 494 TCCCTCCGCGGATCCGACACCTCCCTGGGCTATCTTAAGCTGCCCTTCCATAC 553
QY 653 TACTTGGACTCTTCAAGTGTACGTAGAGGGGCTTGGCTCCGAGATTCCTTAAGTTC 712
Db 554 TACTTGGACTCTTCAAGTGTACGTAGAGGGGCTTGGCTCCGAGATTCCTTAAGTTC 613
QY 713 CTTGGCTCCAGAGACATCCACACCTCCCTACCCACCCGCTCTCCACACCCCTC-G 771
Db 614 GCGGCTCC-----CTTGACAGCTCTCTGGGACACCGGCTCCCTCTGCCCCACCCCTAG 668
QY 772 CTGCTCTTGGTCCAGTCTGTCTCTC--TCAAAGGAGCAGCAGCTTGTACATTT 829
Db 669 CCGCTCTTGGTCCAGTCTGTCTCTC--TCAAAGGAGCAGCAGCTTGTACATTT 728
QY 830 TCCATTC 837
Db 729 TTCCATTC 736

```

```

RESULT 11 834 bp mRNA linear EST 25-SEP-2001
LOCUS B1766766
DEFINITION 60305686F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206217 5',
mRNA sequence.
ACCESSION B1766766
VERSION B1766766.1 GI:15758344
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE NIH-MGC http://mgc.ncl.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: L1M11517 row: c column: 18
High quality sequence stop: 772.
Location/Qualifiers

```

```

FEATURES
Source
1..834
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5206217"
/clone_id="NIH_MGC_122"
/lab_host="DH10B"

```

```

/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC Library."

```

```

BASE COUNT 154 a 287 c 223 g 170 t
ORIGIN

```

```

Query Match 40.5%; Score 473; DB 13; Length 834;
Best Local Similarity 78.0%; Pred. No. 4,4e-112;

```

```

Matches 653; Conservative 0; Mismatches 160; Indels 24; Gaps 6;
QY 96 CTCAGAGAGACTGACACAGACAGACCGCGGAGCCCTCGAACTGAATCCCAAGAC 155
Db 1 CCCAGAGAGACTGTGTGCAAGAGAGACAGACCGCTCGAACTGAATCCCAAGAC 60
QY 156 AGGAAGCCAGAGATGTGTACTTCTTGAACAACACTAGTCCGGGCTGGAAGATGTC 215
Db 61 AAGAAAGCCAGAGATCTGTGCTTCTTGAACAACACTAGTCCGGGCTGGAAGATGTC 120
QY 216 CTAAGGCGGAGAGCGGCGCTCGCGAGCTATTCAGCCATATATAGAGTTATCTC 275
Db 121 CTAAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 276 GGCC-AGACAGAGATGAGAGACAAAGAGATGATGGAGAGAGAGAGAGAGAGAG 334
Db 181 GACCGTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 335 ACCAAATCAACAGCTCCAGGCTCTGCGCTAGCAGACCGCAGATTTGGGAATTAAGTC 394
Db 241 GCCAAGATCAACAGCTCCAGGCTCTGCGCTAGCAGACCGCAGATTTGGAGATTAAGTC 300
QY 395 ATCAGGCGTGGCTCTACTACTCTGATGTCAGGTGACCTTGTATGAGAGAGAGTC 454
Db 301 ACCGCGCTGGGCTCTACTACTCTGATGTCAGGTGACCTTGTATGAGAGAGAGTC 360
QY 455 TACCTGAAGCTGAGCTTGTGTGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 514
Db 361 TACCTGAAGCTGAGCTTGTGTGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
QY 515 TCAGCAGACAGAGAGAGAGCTCTGAGGCGCCAGCTCGCTTGTGTGAGAGGTGTGT 574
Db 421 TCAGCAGCTGAGGAGAGAGTCTCTGAGGCGCCAGCTCGCTGTGTGTGTGTGTGT 480
QY 575 TTGCGGCTGGGCGGAGGCTTCTCTCTGATGTCGACACCTCCCTGGGCTATCTTAAG 634
Db 481 TTGCGGCTGGGCGGAGCA-GGTCTCTCTGAGAGAGAGAGAGAGAGAGAGAGAG 539
QY 635 GCTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 694
Db 540 GCTGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 598
QY 695 CAGATTCCTTAAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 750
Db 599 -----CCCAAGAGTGTCTCCAGGCTCGCGGCTCTCTCTCTCTCTCTCTCTCTCT 652
QY 751 CCCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 808
Db 653 CCCCTTCTGCCCCCAACCTCAGCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 712
QY 809 GCCAGAGCTTGTCTACATGTTTCCATTC-----ACAGAGTATCTTGTCTTCT 859
Db 713 GCGTGGGCTTGTCTACATGTTTCCATTC-----ACAGAGTATCTTGTCTTCT 772
QY 860 TAACTCCATCCACACCAACACATTCACACTACACTACCTCCCAAGCCCTACTTAT 916
Db 773 AAAAATTCGCCCAACGCCCTCTCCACCTCATTAAGTCCCAATTCCTTACCTTT 829

```

```

RESULT 12 948 bp mRNA linear EST 16-JUN-2002
LOCUS B0707185
DEFINITION A6ENKOURT-8353983 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6278608
5', mRNA sequence.
ACCESSION B0707185
VERSION B0707185.1 GI:21846084
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/.
AUTHORS

```

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Mark Watson  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

FEATURES  
source 1. 948  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5278608"  
/clone\_1lb="NIH\_MGC\_113"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: spleen; Vector: pOT7; Site:1; XhoI; Site:2;  
EcoRI; cDNA made by oligo-dT priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGAG(G). Library constructed by Ling Hong in the  
Laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC Library."

BASE COUNT 173 a 334 c 252 g 187 t 2 others  
ORIGIN

Query Match 40.1%; Score 467.8; DB 14; Length 948;  
Best Local Similarity 76.2%; Pred. No. 1.1e-110;  
Matches 632; Conservative 0; Mismatches 179; Indels 18; Gaps 4;

QY 129 AGCCCTGTAAGTGAATCCCGACAGAGAAAGCCAGATGTGTGTAACCTTTTGGAC 188  
DB 1 ACCCGTCGGAACTGAATCCCGACAGAGAAAGCCAGATGTGTGTAACCTTTTGGAC 60  
QY 189 AACTAGTCGGGCTCGAAGAAAGTGTCTTAAAGCCGGAAGGCGGCGCTCGCGAGCTA 248  
DB 61 GACTAGTTGGGCTCGAAGAAAGTGTCTTAAAGCCGGAAGGCGGCGCTCGCGAGCTA 120  
QY 249 TTGACCCCATTTATGAGTTGATCTCTCGGCGAGACAGATGAGACCAAGAGAGTGTG 308  
DB 121 TCCGACCCCATTTATGAGTTGATCTCTCGGCGAGACAGATGAGAGAGAGAGAGC 176  
QY 309 ATGGGACAGTGAAGTGTGAGAGAGACCAAAATCAACAGCTCAGCCCTCTGCTAGC 368  
DB 177 ACCGGACAGTGAAGTGTGAGAGAGACCAAAATCAACAGCTCAGCCCTCTGCTAGC 236  
QY 369 ACCGACAGATTTGGGAAATTTACAGTATCAGGCTGGGCTCTACTACTGCTAGTACAG 428  
DB 237 ACCGACAGATTTGGGAAATTTACAGTATCAGGCTGGGCTCTACTACTGCTAGTACAG 236  
QY 429 TGCAGTTGATGAGGAAAGAGTGTCTACTGAGCTGGAAGTGTGAGAGAGTGTG 488  
DB 297 TGCAGTTGATGAGGAAAGAGTGTCTACTGAGCTGGAAGTGTGAGAGAGTGTG 356  
QY 489 TGGCCCTGGGCTCGTGGAGAAATTTCTCAGCCACAGCAGAGTCTCTGGGCCCCAGC 548  
DB 357 TGGCCCTGGGCTCGTGGAGAAATTTCTCAGCCACAGCAGAGTCTCTGGGCCCCAGC 416  
QY 549 TCGGTTTGGCCAGTGTGTGGGCTGTGGGCGGCGAGGCTTCCCTCGGATCC 608  
DB 417 TCGGCTCTGCCAGGTGTGTGGGCTGTGGGCGGCGAGGCTTCCCTCGGATCC 476  
QY 609 GCACCTCCGCTGGCTCATTTAAGCTGCCCCCTTCTTAACCTATTGAGACTTTTC 668  
DB 477 GCACCTCCGCTGGGCTCATTTAAGCTGCCCCCTTCTTAACCTATTGAGACTTTTC 536  
QY 669 AAGTCTCAGAGGCGCTGTCTCCAGATTCCTTAACCTTTCCCTGGCTCC---AGGA 725

DB 537 AGGTCAGTGAAGGCGGCTGTCTCCCGACAGTGTCCAGAGCTGCGGCTCCCTCGAC 596  
QY 726 GATATCCACACACTCCCTTACCCACCCACATCTCTCCACCCCTGCTCTGTGTGTC 785  
DB 597 AGCTCTGGGACACCGGTCCTGTGGCCACACCCCTGCTCTGTGTGTGTCAGACCT 656  
QY 786 AGTCTGTCTCTCTCTCAAGAGCAGCAGATGTTTCAATGTTT-----CAATTCGA 838  
DB 657 GCGCCCTCCCTCT---ANNAGCTGCTGGGCGCTGTTCAGAGTGTTCATCCCTTAA 712  
QY 839 CAGAGCTATCTTGTCTCTTAACTCCATCCACACCAACTATCATCTACTAGCT 898  
DB 713 TACAGATTTCCACACTTTTATTACAAATCCCGCCAGCCAAATTCACACCTATAGTT 772  
QY 899 CCCAAGCCCTTACTATATCCCTGACCTGCGCCACCCACCTCCAGCCGACAC 947  
DB 773 CCCAATTCCTGACCCCTTTGAGGCGCCAGGGAGATTCACCTTCCCTCC 821

RESULT 13  
BI596681 828 bp mRNA linear EST 07-SEP-2001  
LOCUS 603243254F1 NIH\_MGC\_96 Homo sapiens cDNA clone IMAGE:5285892 5',  
DEFINITION mRNA sequence.  
ACCESSION BI596681 GI:15489620  
VERSION BI596681.1  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 828)  
NIH-MGC <http://mgi.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shuraki  
Toshlyukki and Piero Carninci (RIKEN)  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LLML1722 row: k column: 13  
High quality sequence stop: 776.

FEATURES  
source 1. 828  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5285892"  
/clone\_1lb="NIH\_MGC\_96"  
/tissue="type="hypothalamus"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pBluescript (modified  
pBluescript KS+); Site:1; BamHI; Site:2; SalI-XhoI (gtcgag  
); Oligo-dT primed using primer 5'-TTTGTGTTTGTGTTTGTGTTVN-3',  
size selected for average insert size 2.3 kb and  
normalized to 10^5. This is a primary library enriched  
for full-length clones and constructed using the  
Cap-trapper method (Carninci, in preparation). Library  
constructed by M. Brownstein (NIH/NHGRI, National  
Institutes of Health). Note: this is a NIH\_MGC Library."

BASE COUNT 155 a 278 c 223 g 172 t  
ORIGIN

Query Match 39.1%; Score 456.4; DB 13; Length 828;  
Best Local Similarity 82.6%; Pred. No. 9.3e-108;  
Matches 550; Conservative 0; Mismatches 106; Indels 10; Gaps 2;

QY 137 GAATGATATCCCGACAGAGAAAGCAGATGTGTGTAACCTTTTGGAACTAGTAC 196

Db	Accession	Gene	Species	Accession	Gene	Species
Db	44	GAACGAAATCCCGACAGAGAAAGAACGACAGATCCCTGGCCCTTCTCCTAACCGATTAAGT	103			
Qy	197	CGGCTCGAAGAAAGTGTCTCTTAAGGCGGGAAGGCGGCGCTGCGCCAGCTATTGACGCC	256			
Db	104	CGGCTCCACAAAGTGCACCTAAAGGCGGGAACACGCGGCTCGAAGAGCGATTCACGCC	163			
Qy	257	CATTATGAGGTTCAATCCCGGCGAGACAGATGAGACAAAGCAGGATGAGATGGGACA	316			
Db	164	CATTATGAAAGTTTCAATCCACGACCTGAGACAGGAGCGAGCGAGGATGAGACGGGACA	223			
Qy	317	GTCAGTGGCTGGGAAGAACCCAAATTAACAAGCTCCACGCCCTCTCGGTACAGCCGACG	376			
Db	224	TTGAGTGGCTGGGAGAGAACCCAGATTAACAAGCTCCACGCCCTCTCGGTACAGCCGACG	283			
Qy	377	ATTGGGGAATTTTACAGTCATCAGGCTGAGGCTGAGCTTACTACTGTACTGTGACGACATTT	436			
Db	284	ATCGGCGAGATTATATAGTACACCGGCGCTGGGCTCTACTACTGTACTGTGACGACATTT	343			
Qy	437	GATGAGGGAAGAGCTGTCTACTTGAAGCTGAGACTGCTGCTGCTGCTGCTGCTGCTGCTG	496			
Db	344	GATGAGGGAAGAGCTGTCTACTTGAAGCTGAGACTGCTGCTGCTGCTGCTGCTGCTGCTG	403			
Qy	497	CGCTCTCGGAAGAAATTTCTACAGCCACAGCAGACAGCTCTCCGCGGCGCCAGCTCGGTTG	556			
Db	404	CGCTCTCGGAAGAAATTTCTACAGCCACAGCAGCTCTCCGCGGCGCCAGCTCGGCTTC	463			
Qy	557	TGCGAGGATGTCTGGGCTTTGCGCGCTGCGGCGAGGATCTTCCCTTCGAGATCGCACCTTC	616			
Db	464	TGCGAGGATGTCTGGGCTTTGCGCGCTGCGGCGAGGATCTTCCCTTCGAGATCGCACCTTC	523			
Qy	617	CCCTGGGCTCATCTTAAGGCTGCGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	676			
Db	524	CCCTGGGCTCATCTTAAGGCTGCGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	583			
Qy	677	TGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	736			
Db	584	TGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	636			
Qy	737	CTCTGCT	793			
Db	637	GCT	696			
Qy	794	CT				
Db	697	CT				

Email: mousestewatson.wustl.edu  
 This clone is available royalty-free through LNL ; contact the  
 IMAGE Consortium (infoimage.lnl.gov) for further information.  
 MGI:429769  
 Seq primer: -28ml3 rev2 E7 from Amersham  
 High quality sequence stop: 301.  
 Location/Qualifiers  
 1..471  
 /organism="Mus musculus"  
 /strain="BALB/c"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:696209"  
 /clone\_lib="Barstead mouse heart MPLB3"  
 /sex="mixed"  
 /tissue\_type="heart"  
 /dev\_stage="6 weeks"  
 /lab\_host="DHI08"  
 /note="Organ: heart; Vector: pT73D-Pac (Pharmacia) with a  
 modified polylinker; Site\_1: EcoRI; Site\_2: NotI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer 15'  
 TGTACCAATCTGAAGTGGGAGCGCGCCGCTTTTATTTTATTTT  
 3']; double-stranded cDNA was ligated to Eco RI adaptors  
 [CTGATGTCGTCACG], digested with Not I and cloned into  
 the Not I and Eco RI sites of the modified pT73 vector.  
 Library constructed by Bob Barstead."

BASE COUNT 91 a 141 c 128 g 111 t  
 ORIGIN

Query Match 37.5% Score 437.8; DB 9; Length 471;  
 Best Local Similarity 97.0% Pred. No.5.1e-103;  
 Matches 457; Conservative 0; Mismatches 12; Indels 2; Gaps 1;

OY	274	TCGCGCAGGAGAGATGAGACACAAGCAGAGTGATGGAGGACAGTGGCTGGGAGA	333
Db	3	TGCGATCCAGAGATGAGAGACACAAGCAGAGTGATGGAGGACAGTGGCTGGGAGA	62
OY	334	GACCAAAATCAACAGCTCCAGCCCTCTGCGCTAGACGCCGACAGATTGGGGAATTTACAGT	393
Db	63	GACCAAAATCAACAGCTCCAGCCCTCTGCGCTAGACGCCGACAGATTGGGGAATTTACAGT	122
OY	394	CATCAGGCGCTGGGCTCTACTACCTGACTGTGACAGTGACATTTGATGAGGGAAGGCTGT	453
Db	123	CATCAGGCGCTGGGCTCTACTACCTGACTGTGACAGTGACATTTGATGAGGGAAGGCTGT	182
OY	454	CTACCTGAGAGCTGAGACTGTGTGTAAAGGTGTGTGTGGCCCTGCTCTGGAAGATT	513
Db	183	CTACCTGAGAGCTGAGACTGTGTGTAAAGGTGTGTGTGGCCCTGCTCTGGAAGATT	242
OY	514	CTCAGCCACAGCAGCAGAGCTCTCTGAGGCCGCCAGCTCCGTTTGTGACAGGTGTGGGCT	573
Db	243	CTCAGCCACAGCAGCAGAGCTCTCTGAGGCCGCCAGCTCCGTTTGTGACAGGTGTGGGCT	302
OY	574	GTTGGCGGTGGGGCGCAGGAGCTTCCCTTGGATCCGACACCTCCCTGGGCTCATCTTAA	633
Db	303	GTTGGC--GCTGGGAGGAGGTCTTCCCTTGGATCCGACACCTCCCTGGGCTCATCTTAA	360
OY	634	GAGTCGCCCTTCCCTTGAAGCTTCTTGAAGTCACTAGAGGAGGCTCTGCTCTG	693
Db	361	GAGTCGCCCTTCCCTTGAAGCTTCTTGAAGTCACTAGAGGAGGCTCTGCTCTG	420
OY	694	CCAGATTCCTTAAACTTTCCTGGCTCCAGAGACATCACACACCTCCCTA	744
Db	421	CCAGATTCCTTAAACTTTCCTGGCTCCAGAGACATCACACACCTCCCTA	471

RESULT 15  
 LOCUS B0671259 963 bp mRNA linear EST 15-JUL-2002  
 DEFINITION AGENCOURT B303564 NIH\_MGC\_102 Homo sapiens cDNA clone IMAGE:6274716  
 5', mRNA sequence.  
 ACCESSION B0671259  
 VERSION B0671259.1 GI:21782093





GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 07:08:44 ; Search time 267.523 Seconds

(without alignments)  
9832.178 Million cell updates/sec

Title: US-09-245-198A-1

Perfect score: 1168  
Sequence: 1 ggtgctgagcctggcctg.....ataatcatgattctcttc 1168

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

N.GeneSeq\_101002:\*

1: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1168	100.0	1168	AAV18599	Mus musculus tumou
2	699.4	59.9	701	AAAX23425	Mouse TNRL3 DNA
3	628.6	53.8	1353	21 AAA49717	Human PRO207 CDNA
4	628.6	53.8	1353	24 ABR40255	CDNA encoding huma
5	628.6	53.8	1421	20 AAX56000	Human tumour necro
6	618.2	52.9	1364	24 ABR34881	Human CDNA encodin
7	614.6	52.6	1373	19 AAV18600	Human sapiens tumou
8	597.8	51.2	1236	19 AAV47613	TNF related endoch
9	597.8	51.2	1236	22 AAD04350	Human TREPA (TNF r

10	522.8	44.8	1030	20 AAX23424	Human TNRL3 DNA.
11	498.8	42.7	898	22 AAS03964	Expression vector
12	82.8	7.1	282	16 AAT22190	Human gene signatu
13	69.8	6.0	195	24 ABR29540	Human gene signatu
14	65	5.6	65	24 ABR55975	Mouse adenocarcino
15	46.6	4.0	53522	24 ABR55975	Mouse spliced tran
16	46.6	4.0	53522	19 AAD30228	Human PKD1 gene.
17	46.6	4.0	53577	17 AAT94101	Human PKD1 gene.
18	46.6	4.0	53577	19 AAT94108	Human polykystic k
19	45.8	3.9	105325	20 ABR94407	Human PKD1 locus b
20	45	3.9	1337	24 ABR94407	DNA encoding endot
21	44.8	3.8	1000	21 AAZ17263	Human gene express
22	43.6	3.7	16157	21 AAA02484	Human gene express
23	43.6	3.7	16157	24 ABR70254	Human colon cancer
24	43.6	3.7	16157	24 ABR33083	Chemically treated
25	43.2	3.7	16157	24 ABR34529	Human immune syste
26	43.2	3.7	720	22 ABR51826	Human metastasis a
27	43.2	3.7	2260	22 AAH05001	Ser-Pro-Pro-Pr
28	43.2	3.7	2272	22 AAH18456	Human CDNA clone (
29	43.2	3.7	2272	22 AAS34689	Human CDNA sequenc
30	43.2	3.7	2272	22 AAS34689	Human DNA for a no
31	43	3.7	8604	24 ABR71005	Human DNA for a no
32	43	3.7	524	24 ABR34340	Listeria monocytog
33	42.8	3.7	1166	20 AAZ19440	Oligonucleotide fo
34	42.8	3.7	1166	20 AAZ19440	Oligonucleotide fo
35	42.6	3.6	320	21 AAB38183	Oligonucleotide fo
36	42.6	3.6	556	24 ABR36994	Oligonucleotide fo
37	42.6	3.6	556	24 ABR36995	Oligonucleotide fo
38	42.6	3.6	676	24 ABR36768	Oligonucleotide fo
39	42.6	3.6	1286	24 ABR36768	Oligonucleotide fo
40	42.4	3.6	10732	21 AAB10554	Gene encoding a su
41	42.4	3.6	14006	24 ABR33959	Human immune syste
42	41.8	3.6	1164	24 ABR68833	Listeria monocytog
43	41.6	3.5	320	21 AAB38186	Primer used in the
44	41.4	3.5	1235	20 AAZ16147	Human gene express
45	41.4	3.5	1235	20 AAZ16147	Human gene express

#### ALIGNMENTS

RESULT 1	AAV18599 standard; CDNA: 1168 BP.
ID	AAV18599
AC	AAV18599;
XX	21-JUL-1998 (first entry)
DE	Mus musculus tumour necrosis factor related ligand (TRELL) gene.
XX	TRELL; tumour necrosis factor related ligand; tnf; treatment;
KW	cancer; autoimmune disease; immune system; stimulation; suppression;
KW	graft rejection; ds.
XX	
OS	Mus musculus.
XX	
FH	
FT	Key
FT	Location/Qualifiers
FT	2..679
FT	/*tag= a
FT	/note= "tumour necrosis factor related ligand"
XX	
XX	WO9805783-A1.
XX	
XX	12-FEB-1998.
XX	
XX	07-AUG-1997; 97WO-US13945.
XX	
XX	18-MAR-1997; 97US-0040820.
XX	07-OCT-1996; 96US-0023541.
XX	18-OCT-1996; 96US-0028515.
XX	
XX	(BIO) BIOGEN INC.

(UYGE-) UNIV GENEVA FACULTY MEDICINE.  
 PA Browning JL, Chicheportliche Y;  
 XX MPI: 1998-145619/13.  
 XX P-PSDB; AAM47524.  
 DR  
 PT Tumour necrosis factor related ligand - useful for, e.g. treating  
 XX cancer, auto-immune disease and immune responses to tissue grafts  
 PS  
 XX Claim 2; Pages 45-46; 69pp; English.  
 CC The sequence is that encoding mouse tumour necrosis factor related  
 CC ligand (TRELL). TRELL or active fragments can be included with a  
 CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
 CC diseases or immune responses to tissue grafts, or to stimulate or  
 CC suppress the immune system. It is useful to screen for TRELL  
 CC receptors, by labelling with a detectable label and screening  
 CC compositions for binding. Agents interfering with TRELL-receptor  
 CC binding can also be screened for, can then be administered,  
 CC optionally with interferon- gamma, to induce cell death or  
 CC treat, suppress or alter immune responses (especially involving human  
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
 CC receptor. The DNA sequence can be used in gene therapy for  
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
 CC autoimmune and inflammatory diseases or inherited genetic disorders,  
 CC by introducing into cells, and expressing, therapeutically effective  
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
 CC It may also be of use in the preparation of prepare probes for  
 CC screening natural/synthetic DNAs for TRELL-encoding sequences  
 CC and for antisense therapy.  
 XX  
 XX Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other;  
 SQ  
 Query Match 100.0%; Score 1168; DB 19; Length 1168;  
 Best Local Similarity 100.0%; Pred. No. 8.2e-307;  
 Matches 1168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 GCCCAGCTCCGTTTGCCAGAGTGTGGGCTGTGCGCTGCGCGCCAGGGCTTCCT 600  
 DB 541 GCCCAGCTCCGTTTGCCAGAGTGTGGGCTGTGCGCTGCGCGCCAGGGCTTCCT 600  
 QY 601 TCGGATCCGACCCCTCCCTGGGCTCATCTTAAGGTGCCCCCTTCCTACTACTTGG 660  
 DB 601 TCGGATCCGACCCCTCCCTGGGCTCATCTTAAGGTGCCCCCTTCCTACTACTTGG 660  
 QY 661 ACTCTTCAAGTTCATGAGGGGCTTGGCTCCGAGATTCCTTAACCTTCCTGGCTC 720  
 DB 661 ACTCTTCAAGTTCATGAGGGGCTTGGCTCCGAGATTCCTTAACCTTCCTGGCTC 720  
 QY 721 CAGGACATACACACACCTCCCTACCCACCCGACCTCCGACCCCTGCTCTCT 780  
 DB 721 CAGGACATACACACACCTCCCTACCCACCCGACCTCCGACCCCTGCTCTCT 780  
 QY 781 GGTCCAGCT 840  
 DB 781 GGTCCAGCT 840  
 QY 841 GAGTATCT 900  
 DB 841 GAGTATCT 900  
 QY 901 CAAAGCCCTACTTATCTCTGACTCCCAACCCACCTACCCGACAGCTTTATTGACT 960  
 DB 901 CAAAGCCCTACTTATCTCTGACTCCCAACCCACCTACCCGACAGCTTTATTGACT 960  
 QY 961 TTGTGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1020  
 DB 961 TTGTGACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1020  
 QY 1021 TAGGCCAGAAAGTCCCACTGTGAGGGGAGAGAGCTGGGAGCAAGCTCCCTGATCC 1080  
 DB 1021 TAGGCCAGAAAGTCCCACTGTGAGGGGAGAGAGCTGGGAGCAAGCTCCCTGATCC 1080  
 QY 1081 CTGTGATTTTGAAGAAAGTATTTTATTTATTTATTTATTTATTTATTTATTTAT 1140  
 DB 1081 CTGTGATTTTGAAGAAAGTATTTTATTTATTTATTTATTTATTTATTTATTTAT 1140  
 QY 1141 TAAAGAAATTAATCATGATTTCTCTC 1168  
 DB 1141 TAAAGAAATTAATCATGATTTCTCTC 1168  
 RESULT 2  
 AAX23425  
 ID AAX23425 standard; DNA; 701 BP.  
 AC AAX23425;  
 DT 18-JUN-1999 (first entry)  
 DE Mouse TNRL3 DNA.  
 XX  
 XX Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;  
 KW developmental abnormality; gestational abnormality; prostate cancer;  
 KW AP06; AP08; AP09; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW apoptosis; mouse; ss.  
 OS Mus sp.  
 XX  
 XX Key Location/Qualifiers  
 FT CDS 1..636  
 FT /tag- a  
 FT /product- "TNRL3"  
 XX  
 XX W09911791-A2.  
 XX  
 XX PD 11-MAR-1999.  
 XX

PF 04-SEP-1998: 98MO-US18933:  
 XX  
 PR 05-SEP-1997: 97US-0924634.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 PI Chaudhary PM;  
 XX  
 DR WPI: 1999-205191/17.  
 DR P-PSDB: AAW93591.  
 XX  
 PT New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 PS  
 XX Example VII: Flg 13B; 156pp; English.  
 XX  
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TRRL1 and TRRL3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC moiety. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the changer in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active  
 CC fragments and APO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
 CC activity. The method is performed in vivo or in vitro. APO polypeptides  
 CC are all useful as immunogens for preparing antibodies. APO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 CC  
 XX  
 SO Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;

Query Match 59.9%; Score 699.4; DB 20; Length 701;  
 Best Local Similarity 99.9%; Pred. No. 1.1e-179;  
 Matches 700; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 44 CTGCTGCTGCTCAGCTGGGAGCTGGGCAACCTGTCTGCCAGAGACCTTTCAGAG 103  
 DB 1 CTGCTGCTGCTCAGCTGGGAGCTGGGCAACCTGTCTGCCAGAGACCTTTCAGAG 60  
 QY 104 GACCTGACGACAGAGGACCGCGGAGCGCCCGAAGTGAATCCAGACAGAGAAAGC 163  
 DB 61 GACCTGACGACAGAGGACCGCGGAGCGCCCGAAGTGAATCCAGACAGAGAAAGC 120  
 QY 164 CAGAGTGTGTACTTCTTCTTGAACAACACTAGTCCGGCTCGAAGAGTCTCTTAAAGC 223  
 DB 121 CAGAGTGTGTACTTCTTCTTGAACAACACTAGTCCGGCTCGAAGAGTCTCTTAAAGC 180  
 QY 224 CGAAGAGCGCGCTCGCGAGCTATTGCAAGCCCATTTAGAGGTTCATCTCGGACAGA 283  
 DB 181 CGAAGAGCGCGCTCGCGAGCTATTGCAAGCCCATTTAGAGGTTCATCTCGGACAGA 240  
 QY 284 CAGAGTGAACACAGAGAGGTGAGTGGAGAGTGAAGTGGCGGAGAGACAAATC 343  
 DB 241 CAGAGTGAACACAGAGAGGTGAGTGGAGAGTGAAGTGGCGGAGAGACAAATC 300  
 QY 344 AACAGTCCAGCCCTCTGCGCTAGCAGCCGACAGATTGGGGAATTTACATCAGAGGCT 403  
 DB 301 AACAGTCCAGCCCTCTGCGCTAGCAGCCGACAGATTGGGGAATTTACATCAGAGGCT 360  
 QY 404 GGGCTCTACTACTGTACTGTCTCAGGTGACACTTTGATGAGGAGAAAGCTGTCTACTGAAG 463  
 DB 361 GGGCTCTACTACTGTACTGTCTCAGGTGACACTTTGATGAGGAGAAAGCTGTCTACTGAAG 420

QY 464 CTGCACTGTGCTGTAAGCGGTGTCGCGCCCTGCGCTGCGTGAAGAATTTCAGCACA 523  
 DB 421 CTGCACTGTGCTGTAAGCGGTGTCGCGCCCTGCGCTGCGTGAAGAATTTCAGCACA 480  
 QY 524 GCAGCAAGCTCTCTGCGGCGCCAGCTCCGTTTGTCCAGAGTGTCTGGGCTGTGCGCTG 583  
 DB 481 GCAGCAAGCTCTCTGCGGCGCCAGCTCCGTTTGTCCAGAGTGTCTGGGCTGTGCGCTG 540  
 QY 584 CGGCAAGGCTCTTCCCTTGCGGATCCGACCCCTCCCTGGGCTCATCTTAAAGCTGCCCC 643  
 DB 541 CGGCAAGGCTCTTCCCTTGCGGATCCGACCCCTCCCTGGGCTCATCTTAAAGCTGCCCC 600  
 QY 644 TTCTTAACCTACTTGTGAGACTTTCAAGTTCACTGAGGGGCTTGTCTCCAGATTCT 703  
 DB 601 TTCTTAACCTACTTGTGAGACTTTCAAGTTCACTGAGGGGCTTGTCTCCAGATTCT 660  
 QY 704 TAACTTCCCTGCTCAGAGACATCAGACATCAGACATCAGACATCAGACATCAGACAT 744  
 DB 661 TAACTTCCCTGCTCAGAGACATCAGACATCAGACATCAGACATCAGACATCAGACAT 701

RESULT 3  
 AAA49717  
 ID AAA49717 standard; cDNA; 1353 BP.  
 XX  
 AC AAA49717;  
 XX  
 XX  
 DT 25-SEP-2000 (first entry)  
 XX  
 DE Human PRO207 cDNA clone DNA30879-1152.  
 XX  
 KW PRO207; human; antitumor; tumour; therapy; cytostatic;  
 KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
 KW uterine cancer; prostate cancer; lung cancer; bladder cancer;  
 KW central nervous system cancer; melanoma; leukemia; neoplasm; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 58..807  
 FT sig\_peptide 58..177  
 FT mat\_peptide 178..804  
 FT /tag= a  
 FT /tag= b  
 FT /tag= c  
 PN WO200037638-A2.  
 PD 29-JUN-2000.  
 XX  
 PF 02-DEC-1999: 99MO-US28565.  
 XX  
 PR 22-DEC-1998: 98US-0113296.  
 PR 08-MAR-1999: 99MO-US05028.  
 PR 21-APR-1999: 99US-0130232.  
 PR 28-APR-1999: 99US-0131445.  
 PR 14-MAY-1999: 99US-0134287.  
 PR 20-JUL-1999: 99US-0144758.  
 PR 26-JUL-1999: 99US-0145698.  
 PR 15-SEP-1999: 99MO-US21090.  
 PR 15-SEP-1999: 99MO-US21547.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;  
 PI Napier MA, Pletli RM, Wood WI;  
 XX  
 DR WPI: 2000-442668/38.  
 DR P-PSDB: AAY95358.  
 XX  
 PT Novel composition to inhibit neoplastic cell growth or for treating  
 PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,  
 PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or















PN	US6207642-B1.
XX	27-MAR-2001.
PD	
XX	26-JUN-1998; 98US-0105343.
PF	
XX	12-FEB-1997; 97US-0798692.
PR	10-FEB-1998; 98US-0021706.
XX	(ABBO ) ABBOTT LAB.
PA	
XX	Wiley SR:
PI	
XX	WPI: 2001-280760/29.
DR	P-PSDB: AAEO0891.
XX	
PT	Inducing angiogenesis in mammal at desired sites for promoting wound healing, by administering soluble fragment of extracellular domain of tumor necrosis factor related endothelium proliferative agent protein
PT	-
XX	
PS	Example 2; Column 73-74; 53pp; English.
CC	The present invention relates to extracellular signal molecules,
CC	particularly members of tumor necrosis factor (TNF) family molecules designated as TREPA (TNF related endothelium proliferative agent).
CC	Soluble biologically active TREPA are used to treat TREPA-associated diseases, tumours or metastases. TREPA is used for inducing angiogenesis in human for promoting wound healing and for vascularising grafted tissue for successful grafting and to promote tissue grafts.
CC	The present sequence is a cDNA clone ID #69050 encoding human TREPA.
CC	
XX	Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;
XX	
QY	Query Match
Best Local Similarity	51.2%; Score 597.8; DB 22; Length 1236;
Matches	879; Conservative 0; Mismatches 222; Indels 68; Gaps
QY	2 GTGCTGAGCCTTGCGGCGTGGCGCTGGCTTCCTTGGCTCTGTGCTGCACGCCG 61 
Db	73 GCCTGCGGCGCTTGCGGCGTGGCGCTGGCGCTTCCTTGGCTCTGTGCTGCACGTTTG 132 
QY	62 GGGAGCTGGGGAAGCGCTGTGCCCCAGAGAAGCTTCTCAAGAGACTGACACAGAGAGC 121 
Db	133 GGGAGCCGGCATCGCTCTCCGCCAGAGAGCTGCCACAGAGACTGGTGACAAGAGAG 192 
QY	122 CGCCGGAGACCCTCGAAGTAATCCCAGACAGAGAGAAAGCAGAGATGGTATTCTTC 181 
Db	193 GACCAGACCGGTGGGAAGTGAATCCCAAGACAGAAAGAACAGAGATCTCGGCTTTC 252 
QY	182 TTGGAACAACTAGTCCGGCTTCGAAGAAGTGTCTTAAAGCCGGAAGGCGGCGCTGCG 241 
Db	253 CTGAACCGACATGTTGGGCTCGAAGAAGTGCACCTTAAAGCCGGGAAAAACAGGGCTCCA 312 
QY	242 CGAGCTATTGAGAGCCATTATGAGGTTTATCTCCGCGACAGAGATGAGCACAGAA 301 
Db	313 AGAGGATTCGACAGCCATTATGAAATTCATCCAGACACTGACAGAGAGAGCGACAGCA 372 
QY	302 GGTTGTGATGGAGACAGTAGTGCTGGGAAGAGACCAAATAACAGACTCCAGCCCTTG 361 
Db	373 GGTGTGGACGGGAGACGTAGTGCTGGGAGAGAAAGCAAAATCAAGCTCCAGCCCTTG 432 
QY	362 CGCTACGAGCCGACGATTGGGGAAATTTACAGATACAGAGGCGGGGCTCTACTACCTGAC 421 
Db	433 CGCTACAAACCCGACAGATCGGGAGTTTATGACACCCGGGCTGGGCTCTACTACTGTAC 492 
QY	422 TGTCAGGTGCACTTTGATGAGGAAAAGCGCTGTCTACTCTGAAGCTGGACTTCTGTGAA 481 
Db	493 TGTCAGGTGCACTTTGATGAGGAAAAGCGCTGTCTACTCTGAAGCTGGACTTCTGTGAA 552 
QY	482 GGTGTGCTGGGCGCTGGCTGCTGGAAGAATTTCTAGCCACAGACAGCAAGCTCTCTGGG 541 
Db	553 GGTGTGCTGGGCGCTGGCTGCTGGAAGAATTTCTAGCCACAGCTGGCGAGTTCTCTGGG 612 

QY	542	CCCCGAGTCCGTTTGTGGCAGGTGTCTGGGCTGTGGTCCGGTGGCGGACAGGGTCTCCCTT	601
Db	613	CCCCGAGTCCGTTTGTGGCAGGTGTCTGGGCTGTGGTCCGGTGGCGGACAGGGTCTCCCTT	672
QY	602	CGGATCCGACCCCTCCCTGGGCTCATCTTAAGGCTGGCCCTTCTTAACCTACTTTGGA	661
Db	673	CGGATCCGACCCCTCCCTGGGCTCATCTTAAGGCTGGCCCTTCTTAACCTACTTTGGA	732
QY	662	CTCTTTCAAGTTCACTGAGGGGCTGTGCTCTGCCAGATTTCTTAACCTTTCCCTGGCTCC	721
Db	733	CTCTTTCAAGTTCACTGAGGGGCTGTGCTCTGCCAGATTTCTTAACCTTTCCCTGGCTCC	792
QY	722	AGGACATCACACACCTCCCTACCCCAACCCCACTCTCCACCCCTTC- GCTGCTCTT	780
Db	793	AGGACATCACACACCTCCCTACCCCAACCCCACTCTCCACCCCTTC- GCTGCTCTT	847
QY	781	GGTCAGTCTGCTCTCTCC- TCAAAAGCAGACAGAGCTGTTCACATGTTTCATTCC- 837	
Db	848	GGTCAGTCTGCTCTCTCC- TCAAAAGCAGACAGAGCTGTTCACATGTTTCATTCC- 907	
QY	838	-----ACAGACGTATCTTGTCTCTTTAACTATCCATCCACACATATCAGCTC 891	
Db	908	ACATTAATACAGTATTTCCCACTCTATCTTACAAACACCCACCGCCACCTCCACCTC 967	
QY	892	ACTAGCTCCCAAGGCCCTTAC-----TTATCCGTGACTCCCCACCCCACT 936	
Db	968	ACTAGCTCCCAAGGCCCTTAC-----TTATCCGTGACTCCCCACCCCACT 1027	
QY	937	CACCGACACAGTGTTTATTTGACTTTGTCAC----- 968	
Db	1028	CAGACCCCGAGGGCATTTGTGTTCACTGTACCTGTGGGCAAGATGGGTCCAGAAAGCC 1087	
QY	969	-----CAGGCACTGAGATGGGCTGGACCTGGTGGCAGAGAACCACTGGGACTAG 1023	
Db	1088	CACCTCAGGCACTTAAGAGGGGCTGGACCTGGCAGGAGAACCAAGACTGGGCTAG 1147	
QY	1024	GCCAGAGTTCGCAACTGTGAGGGGAGAGAGCTGGGACCAAGCTCCCTGGA-----TC 1079	
Db	1148	GCCAGAGTTCGCAACTGTGAGGGGAGAGAGCTGGGACCAAGCTCCCTGGA-----TC 1206	
QY	1080	CCTGTGATTTTGAAGAAGTACTATTTT 1108	
Db	1207	CCTGTGATTTTGAAGAAGTACTATTTT 1235	
RESULT 10			
AXX	AXX23424		
AXX	AXX23424	standard; DNA; 1030 BP.	
AXX	AXX23424;		
AXX	AXX23424;		
DT	18-JUN-1999	(first entry)	
XX	DE	Human TNRL3 DNA.	
KW	XX	Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;	
KW	XX	developmental abnormality; gestational abnormality; prostate cancer;	
KW	AP06;	AP08; AP03; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;	
KW	cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;		
KW	apoptosis; human; ss.		
XX	OS	Homo sapiens.	
XX	XX		
FT	Key	Location/Qualifiers	
FT	CDS	1..627	
FT	FT	/*tag= a	
FT	FT	/product= "TNRL3"	
XX	PN	W09911791-A2.	
XX	XX		
PD	11-MAR-1999.		

PF 04-SEP-1998; 98WO-US18393.  
 XX  
 XX 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 XX  
 PI Chaudhary PM;  
 XX  
 DR WPI: 1999-205191/17.  
 XX P-PSDB: AAM93590.  
 PR  
 PT New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 XX  
 PS Example VII; Fig 13A. 156pp; English.  
 XX  
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC moiety. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the changer in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active  
 CC fragments and APO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
 CC activity. The method is performed in vivo or in vitro. APO polypeptides  
 CC are all useful as immunogens for preparing antibodies. APO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 XX  
 S0 Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;  
 Query Match 44.8%; Score 522.8; DB 20; Length 1030;  
 Best Local Similarity 81.6%; Pred. No. 1.2e-131;  
 Matches 643; Conservative 0; Mismatches 137; Indels 8; Gaps 3;  
 OY 53 GTCACCTGGGAGCTGGGCAACGCTGCTGCCAGAGACCTCTTCAGAGAGAGCTGACA 112  
 DB 1 GTCACCTGGGAGCTGGGCAACGCTGCTGCCAGAGACCTCTTCAGAGAGAGCTGAGT 60  
 OY 113 GCAGAGAGACCGCGGAGACCCCTGAACTGAATCCCGAGAGAGAGAGAGATG 172  
 DB 61 GCAGAGAGAGACCGCGGAGACCCCTGAACTGAATCCCGAGAGAGAGAGATGCT 120  
 OY 173 GTCACCTGGGAGCTGGGCAACGCTGCTGCCAGAGAGCTCTTAAAGCGCGAAGCG 232  
 DB 121 GCGCCTTTCCTGAACCGCTAGTGGCTCCAGAGAGCTCACTAAAGCGCGAAGAA 180  
 OY 233 GCGCCTGGCGGATGATGACCGCATATATAGTTCATCCCGGCGAGAGAGATGGA 292  
 DB 181 GCGGCTCGAAGAGCATGCGACGCCCATATATGAAGTTCAATCCAGCATGACAGAG 240  
 OY 293 GCACAGAGAGAGTGTGATGAGACAGTGAAGTGGTGGAGAGAGACAAAATCAACAGCTCC 352  
 DB 241 GCGAGAGAGAGTGTGATGAGAGAGAGTGAAGTGGTGGAGAGAGAGATCAACAGCTCC 300  
 OY 353 AGCCCTCTGCGCTACGACCGCGAGATGGGGAATTTACAGTATCAGGCGTGGCTTAC 412  
 DB 301 AGCCCTCTGCGCTACGACCGCGAGATGGGGAATTTATAGTACCGCGGCTGGCTTAC 360  
 OY 413 TACCTGTACTGTACAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 472  
 DB 361 TACCTGTACTGTACAGTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420

OY 473 CTGGTGAACGGTGTGCTGGCCCTGGCGTGCCTGGAAGAATTCACAGACAGACAGC 532  
 DB 421 CTGGTGAACGGTGTGCTGGCCCTGGCGTGCCTGGAAGAATTCACAGACAGACAGC 480  
 OY 533 TCTCTGGGCCCCAGACCTCCGTTTGTGCCAGGTGTGCGGCTGTGCGCGGCGCAGG 592  
 DB 481 TCCCTGGGCCCCAGACCTCCGCTTGTGCCAGGTGTGCGGCTGTGCGCGGCGCAGG 540  
 OY 593 TCTTCCCTGGATCCGACACCTCCCTGGGCTCATCTTAAAGGCTGCCCTCTCTAAC 652  
 DB 541 TCCCTCCGCGGATCCGACACCTCCCTGGGCTCATCTTAAAGGCTGCCCTCTCTAAC 600  
 OY 653 TACTTGGACCTTTCAGATTTCACGTAGAGAGGCGCTGTCTCCACAGATCTTAAACTTTC 712  
 DB 601 TACTTGGACCTTTCAGATTTCACGTAGAGAGGCGCTGTCTCCACAGATCTTCCAGGCT 660  
 OY 713 CTTGCTCCAGAGACATACACACCTCTTACCCACCCCTCTCTCCACACCCCTC-G 771  
 DB 661 GCGGCTCC-----CCTGACAGACCTCTTGGGACCCGCTCCCTCTCCACCCCTCAG 715  
 OY 772 CTGCTCCCTGGTCCAGTCTGCTCTCTCC--TCAAGGAGCAGCAGCTTGTTCACATGT 829  
 DB 716 CCGCTCTTGTCCAGACCTGCTCCCTCTAGAGGCTGCTGGGCTGCTTACAGTGT 775  
 OY 830 TCCATTCC 837  
 DB 776 TTCCATTC 783  
 RESULT 11  
 AAS03964  
 ID AAS03964 standard; DNA; 898 BP.  
 XX  
 XX AAS03964;  
 AC  
 XX 26-SEP-2001 (first entry)  
 DT  
 XX  
 DE Expression vector pDC409-L2-TWEAK fusion protein-encoding DNA.  
 XX  
 KW TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;  
 KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 KW rubeosis; uveitis; macular degeneration; arthritis; rheumatism; ds;  
 KW corneal graft neovascularisation; psoriasis; metastatic condition;  
 KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
 KW preneoplastic condition; myocardial angiogenesis; wound granulation;  
 KW scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
 KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 KW peripheral atherosclerosis; pDC409-L2-TWEAK; TWEAK receptor; TWEAKR;  
 KW fusion protein.  
 OS  
 OS Homo sapiens.  
 XX  
 XX Synthetic.  
 FH  
 FT Key Location/Qualifiers  
 FT CDS 52..873  
 FT  
 FT /tag= a  
 FT /product= "Fusion protein comprising a growth hormone  
 FT leader, a leucine zipper multimerisation  
 FT domain, and human TWEAK extracellular  
 FT domain"  
 XX  
 XX MO200145730-A2.  
 XX  
 XX 28-JUN-2001.  
 PD  
 XX 19-DEC-2000; 2000WO-US34755.  
 XX  
 XX 20-DEC-1999; 99US-0172878.  
 PR 10-MAY-2000; 2000US-0203347.  
 XX  
 XX (IMMUNEX ) IMMUNEX CORP.  
 XX

PI WILEY SR;  
 XX  
 DR WPI: 2001-417975/44.  
 DR P-PSDB: AAU03499.  
 XX  
 PT Modulating angiogenesis in a mammal for treating diseases mediated by  
 PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  
 PT peripheral tissue, by administering antagonist or agonist of TWEAK  
 XX  
 PS Example 1; Page 39-40; 46pp; English.  
 XX  
 CC The sequence represents a DNA from the expression vector  
 CC PDC409-12-TWEAK, which encodes a fusion protein comprising a growth  
 CC hormone leader, a leucine zipper multimerisation domain, and the  
 CC extracellular domain of human TWEAK. The fusion protein was used in  
 CC the isolation of human TWEAK receptor (TWEAKR)-expressing clones  
 CC from a COS cell human cDNA library. The TWEAK protein is a  
 CC member of the tumour necrosis factor (TNF) family and induces  
 CC angiogenesis. TWEAKR may therefore be used to screen for and develop  
 CC TWEAKR agonists and antagonists for the modulation of angiogenesis, to be  
 CC used in the treatment and diagnosis of human disease. The disorders  
 CC mediated by angiogenesis include ocular disorders characterised by ocular  
 CC neovascularisation such as diabetic retinopathy, neovascular glaucoma,  
 CC retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,  
 CC rubellosis, uveitis, macular degeneration and corneal graft  
 CC neovascularisation, and inflammatory diseases such as arthritis,  
 CC rheumatism and psoriasis. Other treatable diseases include malignant and  
 CC metastatic conditions such as sarcomas and carcinomas, benign tumours and  
 CC preneoplastic conditions, myocardial angiogenesis, plaque formation and  
 CC scleroderma, vascular adhesions, atherosclerotic plaque  
 CC neovascularisation, telangiectasia, wound granulation, coronary  
 CC atherosclerosis, peripheral atherosclerosis and ischemia.  
 CC  
 XX Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other;  
 SQ  
 Query Match 42.7%; Score 498.8; DB 22; Length 898;  
 Best Local Similarity 87.0%; Pred. No. 3.7e-125;  
 Matches 548; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 56 AGCCGGGAGCTGGGACGACGCTGTCTGCCAGAGAGCTTTCAGAGAGCTGACACCA 115  
 DB 250 AGTTGGGGAGCCGGGCGCTGCTGCCAGAGAGCTTCCAGAGAGCTGAGTGGCA 309  
 QY 116 GAGGACCCCGGAGAGCCCTGAACTGAATCCAGAGAGGAAAGCCAGAGTGTGTA 175  
 DB 310 GAGAGAGACAGAGACCCCTCGAACTGAATCCAGAGAGGAAAGCCAGAGTCTCG 369  
 QY 176 CCTTCTTGAACAACCTAGTCCGGCTTCGAAAGTCTCTTAAAGCCGGAAGCGCGG 235  
 DB 370 CCTTCTTGAACAACCTAGTCCGGCTTCGAAAGTCTCTTAAAGCCGGAAGCGCGG 429  
 QY 236 CCTGCGCAGCTATTTGACAGCCCATTTAGAGTTTATCTCGGCGACAGAGATGAGCA 295  
 DB 430 GCTGGAAGAGGATCGACCCCATTTAGAACTTCATCCAGACCTGACAGAGCGAGCG 489  
 QY 296 CAACAGAGTGTGATGAGACAGTGTGCTGGGAGGAGACCAATTAACAGCTCCAGC 355  
 DB 490 CAGGACAGTGTGAGCGGACAGTGTGCTGGGAGGAGGAGCAATTAACAGCTCCAGC 549  
 QY 356 CCTGTGCGCTAGACCGCCAGATTTGGGAATTTACAGTCAAGCTGGGCTTACTAC 415  
 DB 550 CCTGTGCGCTAGACCGCCAGATTTGGGAATTTATAGTCAACCCGGGCTTACTAC 609  
 QY 416 CTGACGTGATGAGTCACTTTGATGAGGAAAGGCTGTCTTACCTGAAGCTGAGCTTG 475  
 DB 610 CTGACGTGATGAGTCACTTTGATGAGGAGGAGGCTGTCTTACCTGAAGCTGAGCTTG 669  
 QY 476 GTGAACGATGTGCTGGGCGCTGCTGGAGAAATTTACAGCAACAGCAAGCTCT 535  
 DB 670 GTGATGTGTGCTGGGCGCTGCTGGAGAAATTTACAGCAACAGCAAGCTCT 729  
 QY 536 CTGTGGGCCAGGCTCCGTTTGTGCCAAGTGTGTGGGCTGTGTGCCGCTGGCGGCAAGGCTCT 595

DB 730 CTCGGGCCCCACACTCGCTCTGCCAGAGTGTGGGCTGTGGCCCTGGCGCCAGAGTCC 789  
 QY 596 TCCCTGCGATCCGACACCTCCCGGCGTCAATTAAGCGTCCCTCTTACACTAC 655  
 DB 790 TCCCTGCGATCCGACACCTCCCGGCGTCAATTAAGCGTCCCTCTTACACTAC 849  
 QY 656 TTTGGACTCTTTCAGATTCACACTAGAGGGCC 685  
 DB 850 TTCGACTCTTCCAGGTTCTACTAGCGGGCC 879  
 RESULF 12  
 AAT22190  
 ID AAT22190 standard; cDNA to mRNA; 282 BP.  
 AC AAT22190;  
 XX 27-AUG-1996 (first entry)  
 DT  
 DE Human gene signature HUMGS03761.  
 XX  
 KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;  
 KW human; cloning; mapping; non-biased library; diagnosis; detection;  
 KW cell typing; abnormal cell function; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W09514772-A1.  
 XX  
 PD 01-JUN-1995.  
 XX  
 PF 11-NOV-1994; 94WO-JP01916.  
 XX  
 PR 12-NOV-1993; 93JP-0355504.  
 XX  
 PA (MATS/) MATSUBARA K.  
 PA (ORDB/) OKUBO K.  
 PI Matsubara K, Okubo K;  
 PI  
 DR WPI: 1995-206931/27.  
 XX  
 PT Identifying gene signatures in 3'-directed human cDNA library - e.g.  
 PT for diagnosis of abnormal cell function, by preparing cDNA that  
 PT reflects relative abundance of corresp. mRNA in specific human  
 PT tissues  
 PS Claim 1; Page 1067; 2245pp; Japanese.  
 XX  
 CC A single-stranded DNA (or its complementary strand or the corresp.  
 CC double-stranded DNA) which comprises one of the 7837 "GS" sequences  
 CC given in AAT19001-T26837 and which is able to hybridise to part of  
 CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)  
 CC sequences were obtained from 3'-directed cDNA libraries prepared  
 CC from various human tissues; synthesis of cDNA was initiated from the  
 CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-  
 CC untranslated sequence is unique to a particular mRNA species, almost  
 CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library  
 CC is constructed so as to reflect accurately the relative abundance of  
 CC different mRNAs in the particular tissue from which it was derived.  
 CC The appearance frequency of a given GS in a cDNA library can be  
 CC determined (esp. using primers and probes derived from the GS  
 CC sequences) as a means of diagnosing abnormal cell function or for  
 CC recognising different cell types.  
 XX  
 SQ Sequence 282 BP; 80 A; 62 C; 69 G; 66 T; 5 other;  
 Query Match 7.1%; Score 82.8; DB 16; Length 282;  
 Best Local Similarity 80.6%; Pred. No. 1.9e-12;  
 Matches 158; Conservative 0; Mismatches 26; Indels 12; Gaps 5;

QY 969 CAGGCACTGAGATGGGCTGAGCTG-GTGGCAGGAAGCCAGAGACCTGGAGTACGCCA 1027

```

Db      88 CAGGACTAAGAGGGGCTGGACCTNTGGCGGAGGAGCAAGCAAGAGCTGGCGCTAGGGCA 147
      1028 GAAGTCCCAACTGTGAGGGGGAAGAGCTGGGACACAGCTCTCCCTGGA----TCCTGT 1083
      148 GGAGTCCCAAAATGTGAGGGGCGAGA-AACAAGACAGAGCTCTCCCTGGAAGTCCCTGT 206
      1084 TGGATTTTGAAG--AGATACATATTTTATATATATGTCGACAAATGTT----AAATGGA 1137
      207 TGGATTTTAAACAGATATATATTTTNTATATGTGACAAATGTTGTAATATGGA 266
      1138 TATTAAGAGAAATAAA 1153
      267 TATTAATATAGATAAA 282

RESULT 13
ABK29540
ID      ABK29540 standard; cDNA; 195 BP.
XX
AC      ABK29540;
XX
DT      23-APR-2002 (first entry)
XX
DE      Colon adenocarcinoma-specific cDNA #66.
XX
KW      Human; colon adenocarcinoma; colon cancer; tumour; gene; ss.
XX
OS      Homo sapiens.
XX
PN      WO200196389-A2.
XX
PD      20-DEC-2001.
XX
PF      07-JUN-2001; 2001WO-US18574.
XX
PR      09-JUN-2000; 2000US-210667P.
XX
PR      22-NOV-2000; 2000US-252614P.
XX
PA      (CORI-) CORIXA CORP.
XX
PI      Meagher MJ, King GE, Xu J, Secrist H;
XX
DR      WPI; 2002-098052/13.
XX
PT      New isolated polynucleotide encoding a polypeptide comprising a portion
PT      of colon tumour protein, for detection, diagnosis and therapy of human
PT      colon cancer.
XX
PS      Claim 1; Page 133; 211pp; English.
XX
CC      The invention relates to an isolated polynucleotide (I) encoding a
CC      polypeptide (II) comprising a portion of a colon tumour protein. A new
CC      oligonucleotide (III) that hybridises to (I) is useful for
CC      determining the presence of a cancer in a patient. (II) or antigen
CC      presenting cells expressing (I) are useful for stimulating and/or
CC      expanding T cells specific for a tumour protein, by contacting T cells
CC      with (I), (II) or antigen-presenting cells that express (I), (I), (II),
CC      or antigen presenting cells that express (II) are useful for treating
CC      colon cancer in a patient by incubating CD4+ and/or CD8+ T cells isolated
CC      from a patient with (I), (II), or antigen presenting cells that express
CC      (II), so that T cells proliferate, and administering to the patient an
CC      effective amount of the proliferated T cells, thus inhibiting the
CC      development of a cancer in the patient. A new composition is useful for
CC      stimulating an immune response in a patient. (I) or (II) is useful in
CC      vaccines and pharmaceutical compositions for prevention and treatment of
CC      colon cancer and for the diagnosis and monitoring of the cancers. (I),
CC      (II) or an antibody against (II) is useful for detection, diagnosis and/
CC      or therapy of human colon cancer. (I) is useful as a probe or primer for
CC      nucleic acid hybridisation, and in the design and preparation of ribozyme
CC      molecules for inhibiting expression of (II) in tumour cells. ABK29475-
CC      ABK29851 represent human colon adenocarcinoma-specific cDNA sequences of
CC      the invention.
```

```

XX
SO      Sequence 195 BP; 49 A; 51 C; 58 G; 37 T; 0 other;
XX
Query Match      6.0%; Score 69.8; DB 24; Length 195;
Best Local Similarity 83.5%; Pred. No. 5.4e-09;
Matches 91; Conservative 0; Mismatches 17; Indels 1; Gaps 1;
XX
      969 CAGGACTAAGAGGGGCTGGACCTGTGCGGAGGAAAGCCAGAGACCTGGGACTAGGGCAG 1028
      69 CAGGACTAAGAGGGGCTGGACCTGTGCGGAGGAAAGCCAGAGACCTGGGCTAGGGCAG-128
      1029 AAGTCCCAACTGTGAGGGGGAAGAGCTGGGACACAGCTCTCCCTGGA 1077
      129 GAGTTCCTCAATGTGAGGGGCGAGA-AACAAGACAGAGCTCTCCCTTGA 176

RESULT 14
ABN55975
ID      ABN55975 standard; DNA; 65 BP.
XX
AC      ABN55975;
XX
DT      15-JUL-2002 (first entry)
XX
DE      Mouse spliced transcript detection oligonucleotide SEQ ID NO:28723.
XX
KW      Human; mouse; rat; splice transcript; detection; RNA transcript;
KW      splice variant; transcriptome; oligonucleotide library; ss.
XX
OS      Mus musculus.
XX
PN      WO200210449-A2.
XX
PD      07-FEB-2002.
XX
PF      20-JUL-2001; 2001WO-IB01903.
XX
PR      28-JUL-2000; 2000US-221607P.
XX
PR      02-MAY-2001; 2001US-287724P.
XX
PA      (COMP-) COMPUGEN INC.
XX
PI      Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR      WPI; 2002-257383/30.
XX
PT      New oligonucleotide libraries comprising oligonucleotides which
PT      selectively hybridize to mRNAs transcribed from a transcription unit of
PT      a genome, useful for detecting tissue-, pathology-, and
PT      developmental-specific genes.
XX
PS      Example 1; SEQ ID 28723; 47pp; English.
XX
CC      The present invention describes oligonucleotide libraries for detecting
CC      messenger RNAs that populate a (sub-)transcriptome, where the
CC      (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC      transcription units that populate a genome. The library comprises
CC      several oligonucleotides, each capable of hybridising selectively to a
CC      set of messenger RNAs transcribed from a given transcription unit of
CC      the genome, which encodes one or more messenger RNA splice variants.
CC      The oligonucleotide libraries are useful for detecting mRNAs from a
CC      biological sample, in expression profiling studies, in qualitatively or
CC      quantitatively characterising the corresponding transcriptome, and in
CC      detecting RNA transcripts and splice variants of human or animal
CC      transcriptomes. The libraries may also be used as specialised mini
CC      libraries to detect transcripts of a sub-transcriptome under a
CC      particular biological or pathological state, and so allowing the
CC      detection of tissue- and pathology-specific genes such as those genes
CC      only expressed in specific tissue under a specific pathological
CC      condition; to detect developmental specific genes; and to detect RNA
CC      transcripts and splice variants of a transcriptome of a patient suffering
CC      from a particular disorder. ABN27253 to ABN59589 represent
CC      oligonucleotide sequences from rats, humans and mice, which are used in
```

CC the exemplification of the present invention.  
CC N.B. The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 65 BP; 14 A; 15 C; 21 G; 15 T; 0 other;  
Query Match 5.6%; Score 65; DB 24; Length 65;  
Best Local Similarity 100.0%; Pred. No. 6.6e-08;  
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1030 AGTTCACATGTGAGGGGGAAGAGCTGGGACAGACTCCTCCCTGGATCCTGTGATT 1089  
DB 1 AGTTCACATGTGAGGGGGAAGAGCTGGGACAGACTCCTCCCTGGATCCTGTGATT 60  
QY 1090 TTGAA 1094  
DB 61 TTGAA 65  
RESULT 15  
AAD30228  
ID AAD30228 standard; DNA; 53522 BP.  
XX  
AC AAD30228;  
XX  
DT 17-MAY-2002 (first entry)  
XX  
DE Human PKD1 gene.  
XX  
KW Human; PKD1 gene; autosomal dominant polycystic kidney disease; ADPKD;  
KW acquired cystic disease; transgenic animal; chromosome 16; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200206529-A2.  
XX  
PD 24-JAN-2002.  
XX  
PF 13-JUL-2001; 2001WO-US22035.  
XX  
PR 13-JUL-2000; 2000US-218261P.  
XX 13-APR-2001; 2001US-283691P.  
XX  
PA (UY30 ) UNIV JOHNS HOPKINS SCHOOL MEDICINE.  
XX  
PI Germlino GG, Watnick TJ, Phakdeekitchareon B;  
XX  
DR WPI; 2002-179805/23.  
XX  
PT Novel primer for diagnosing polycystic kidney disease-associated  
PT disorder, comprises regions having sequence that selectively hybridizes  
PT to polycystic kidney disease gene sequence -  
XX  
PS Claim 20; Page 127-156; 192pp; English.  
XX  
XX The present invention relates to compositions and methods useful for the  
CC identification and detection of polycystic kidney disease (PKD1) gene  
CC mutations. The invention also relates to primers comprising a 5' region  
CC having a sequence that selectively hybridizes to a PKD1 gene sequence  
CC and optionally, to a PKD1 homologue sequence and an adjacent 3' region  
CC having a sequence that selectively hybridizes to a PKD1 gene sequence  
CC and not to a PKD1 homologue sequence. Primer pairs of the invention are  
CC useful for detecting the presence or absence of a mutation in a PKD1  
CC polynucleotide in a sample, for identifying a subject at risk for a  
CC PKD1-associated disorder such as autosomal dominant polycystic kidney  
CC disease (ADPKD) or acquired cystic disease and for diagnosing a PKD1-  
CC associated disorder in a subject. They are useful for selectively  
CC amplifying a region of a PKD1 gene. PKD1 DNA fragments are useful  
CC detecting the presence of a mutant PKD1 polynucleotide in a sample,  
CC as a probe for an amplification reaction. In hybridisation or  
CC amplification assays of biological samples to detect abnormalities  
CC of PKD1 expression and for engineering transgenic animals. The present

CC sequence is human PKD1 gene located on chromosome 16.  
XX  
SQ Sequence 53522 BP; 8486 A; 17664 C; 15766 G; 11606 T; 0 other;  
Query Match 4.0%; Score 46.6; DB 24; Length 53522;  
Best Local Similarity 48.7%; Pred. No. 0.15;  
Matches 127; Conservative 0; Mismatches 134; Indels 0; Gaps 0;  
QY 685 CTGCTCTCTCCAGATTCTTAACCTTCCCTGGCTCCAGAGCATGCACAGACCTCCCTA 744  
DB 34365 CTTCCCTCTCCCTTACCCCTTTCCCTCTCCCTCCCTAGACCTTCCCTCAGCTCTCC 34424  
QY 745 CCCCACCCCACTCTCCACACCCCTGCTGCTCTTGTCCAGTCTGTCCTCAAA 804  
DB 34425 CGCTGAGCCCTCCACATCGTCCGACCCCTCCCTCCCTAGCCCTCCCTCCCTT 34484  
QY 805 GGCAGCCAGAGCTTGTTCATGTTCATTCACAGAGTATCTGCTCTTAAACA 864  
DB 34485 CCTCCCTCTCCCTCCCTCCCTCCCTCCCTCCCTCTCTCTCCCTCCCTCCCTCC 34544  
QY 865 TCCCATCCACACACACTATACACTACACTAGCTCCCAAGCCCTACTATCCCTGACT 924  
DB 34545 TTCTTCCCTCTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 34604  
QY 925 CCCCACCCCACTGACCCGACC 945  
DB 34605 TCCTCCCTCCCTCTCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 34625

Search completed: May 8, 2003, 07:30:14  
Job time : 314.523 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 8, 2003, 07:13:00 ; Search time 2992.86 Seconds

(without alignments)  
11357.738 Million cell updates/sec

Title: US-09-245-198a-1

Perfect score: 1168

Sequence: 1 ggtgctgagcctggcctgg.....ataaatcatgtctcttc 1168

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_scs:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_din:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1168	100.0	1168	10 AF030100	AF030100 Mus muscu
2	711	60.9	203083	2 AC069459	AC069459 Mus muscu
3	711	60.9	234182	10 AL603707	AL603707 Mouse DNA
4	628.6	53.8	1353	6 AX201324	AX201324 Sequence
5	628.6	53.8	1368	9 AF055872	AF055872 Homo sapi
6	624	53.4	1306	9 AF030099	AF030099 Homo sapi
7	597.8	51.2	1236	6 AR140407	AR140407 Sequence
8	501	42.9	138792	6 AC119115	AC119115 Rattus no
9	498.8	42.7	898	6 BC019047	BC019047 Homo sapi
10	409.4	35.1	1651	9 BC019047	BC019047 Homo sapi
11	304	26.0	215795	2 AC127470	AC127470 Pan trogl
12	303.8	26.0	177703	2 AC016876	AC016876 Homo sapi
13	278.2	23.8	153553	2 AC126921	AC126921 Bos tauru
14	237.4	20.3	177555	2 AC130192	AC130192 Sus scrof
15	212	18.2	161428	2 AC126925	AC126925 Canis fam
16	156.8	13.4	184026	2 AC098923	AC098923 Rattus no
17	111.6	9.6	203281	2 AC126237	AC126237 Canis fam
18	88.4	7.6	7218	6 I66494	I66494 Sequence 14
19	86.4	7.4	139405	2 AC126239	AC126239 Fells cat
20	69.8	6.0	195	6 AX379024	AX379024 Sequence
21	61	5.2	190376	2 AC123372	AC123372 Rattus no
22	58.4	5.0	64424	2 AC124103	AC124103 Mus muscu
23	56.2	4.8	114260	2 AC098956	AC098956 Rattus no
24	56.2	4.8	125020	9 AF429315	AF429315 Homo sapi
25	55.8	4.8	80112	2 AC128800	AC128800 Rattus no
26	35.8	4.8	303091	2 AC084799	AC084799 Mus muscu
27	35.4	4.7	69119	2 AC128983	AC128983 Rattus no
28	55.4	4.7	241570	2 AC094065	AC094065 Rattus no
29	55	4.7	185822	2 AC073554	AC073554 Homo sapi
30	54.8	4.7	100511	2 AC010774	AC010774 Homo sapi
31	54.6	4.7	221758	2 AC068947	AC068947 Mus muscu
32	54.4	4.7	87120	2 AC012225	AC012225 Homo sapi
33	54.2	4.6	298166	2 AC087563	AC087563 Homo sapi
34	54	4.6	111627	2 AC108626	AC108626 Rattus no
35	54	4.6	184036	2 AC126109	AC126109 Rattus no
36	54	4.6	310371	2 AC096296	AC096296 Rattus no
37	53.8	4.6	99517	2 AC131402	AC131402 Rattus no
38	53.6	4.6	205350	2 AC078946	AC078946 Mus muscu
39	53.4	4.6	146986	2 AC095995	AC095995 Rattus no
40	53.4	4.6	203281	2 AC126237	AC126237 Canis fam
41	53.2	4.6	49430	2 AC100434	AC100434 Mus muscu
42	53.2	4.6	182269	2 AC107416	AC107416 Rattus no
43	53	4.5	55061	2 AC091597	AC091597 Mus muscu
44	53	4.5	176645	2 AC114115	AC114115 Rattus no
45	53	4.5	187252	10 AL607109	AL607109 Mouse DNA

#### ALIGNMENTS

RESULT 1  
AF030100 1168 bp mRNA linear ROD 20-DEC-1997  
LOCUS Mus musculus TWEAK mRNA, partial cds.  
DEFINITION AF030100  
ACCESSION AF030100  
VERSION AF030100.1 GI:2707220  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1168)  
REFERENCE  
Hession,C., Garcia,I., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.,  
TWEAK, a new secreted ligand in the tumor necrosis factor family





## JOURNAL

Submitted (31-MAY-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jun 25, 2001 this sequence version replaced gi:12621364.

## COMMENT

Center: Baylor College of Medicine  
Genome Center  
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu

Project Information  
Center project name: RP23-168P5  
Center clone name: RP23-168P5

Summary Statistics  
Sequencing vector: M13: L08821

Chemistry: Dye-terminator Big Dye: 528 of reads  
Assembly program: Phrap: version 0.990329  
Consensus quality: 212648 bases at least Q40  
Consensus quality: 218902 bases at least Q30  
Estimated insert size: 222384 bases at least Q20  
Quality coverage: 0x in Q20 bases; agarose-tp estimation  
Quality coverage: 7.2x in Q20 bases; sum-of-coverage estimation

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 62152: contig of 62152 bp in length  
\* 62153 62252: gap of unknown length  
\* 62253 118772: contig of 56520 bp in length  
\* 118773 118872: gap of unknown length  
\* 118873 148924: contig of 30052 bp in length  
\* 148925 149024: gap of unknown length  
\* 149025 167231: contig of 18207 bp in length  
\* 167232 167331: gap of unknown length  
\* 167332 189907: contig of 22576 bp in length  
\* 189908 190007: gap of unknown length  
\* 190008 196537: contig of 6530 bp in length  
\* 196538 196637: gap of unknown length  
\* 196638 203083: contig of 6446 bp in length.  
Location/Qualifiers:  
1. 203083

FEATURES  
source  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/chromosome="11"  
/clone="RP23-168P5"

BASE COUNT 52662 a 49293 c 47892 g 52633 t 603 others  
ORIGIN

Query Match 60.9%; Score 711; DB 2; Length 203083;  
Best Local Similarity 99.6%; Pired. No. 2.5e-176;  
Matches 744; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

QY 425 CAGGTGACCTTTGATGAGGAAGGCTGTACTGAGCTGAGCTGTGCTGGAAGGCT 484  
DB 43019 CAGGTGACCTTTGATGAGGAAGGCTGTACTGAGCTGAGCTGTGCTGGAAGGCT 42960  
QY 485 GTGCTGACCTGCTGCTGCTGGAAGATTCAGCCACAGCAGCAAGCTCTCTGGGCC 544  
DB 42959 GTGCTGACCTGCTGCTGCTGGAAGATTCAGCCACAGCAGCAAGCTCTCTGGGCC 42900  
QY 545 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604  
DB 42899 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42840  
QY 605 ATCCGACACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 664

DB 42839 ATCCGACACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42780  
QY 665 TTTCAGATTCACAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 724  
DB 42779 TTTCAGATTCACAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42720  
QY 725 AGCATCACACACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 784  
DB 42719 AGCATCACACACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 42660  
QY 785 CAGTCTG 842  
DB 42659 CAGTCTG 42600  
QY 843 CGTATCTG 901  
DB 42599 CGTATCTG 42540  
QY 902 AAGCCCTG 961  
DB 42539 AAGCCCTG 42480  
QY 962 TGTGACACG 1021  
DB 42479 TGTGACACG 42420  
QY 1022 AGCCGACG 1081  
DB 42419 AGCCGACG 42360  
QY 1082 TGTGATTTTGAAGAT 1141  
DB 42359 TGTGATTTTGAAGAT 42300  
QY 1142 AAAGAGATTAATCATGATTTCTCTTC 1168  
DB 42299 AAAGAGATTAATCATGATTTCTCTTC 42273

RESULT 3  
AL603707/c  
LOCUS  
DEFINITION  
VERSION  
ACCESSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus.  
house mouse.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 234182)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT  
Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
On Nov 20, 2001 this sequence version replaced gi:16605765.  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submissions  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=

30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. The following  
abbreviations are used to associate primary accession numbers given



Db	670	CCCCAGCTCCGCTTGGCAGAGTGTGCGGGGCTTTGGGCCATGGGCCAGAGGCTCCCTCG	729
QY	602	CGGATCCGACACCTTCCCTGGGCTCATTTAAGCTGCGCCCTTCTTAACCTACTTTGGA	661
Db	730	CGGATCCGACACCTTCCCTGGGCTCATTTAAGCTGCGCCCTTCTTAACCTACTTTGGA	789
QY	662	CTCTTCAAGTTCATCGTAGGGGCGCTGGCTCCCGAGATTCCTTAACCTTTCCCTGGGCTC	721
Db	790	CTCTTCAAGTTCATCGTAGGGGCGCTGGCTCCCGAGATTCCTTAACCTTTCCCTGGGCTC	849
QY	722	AGGAGCATCACACACTGCTCCCTACCCCAACCCCACTCTCCACCCCTC-GCTGCTCTT	780
Db	850	-----CGTCGACAGCTCTCTGGGACCGGGTCCCTCTGCCCCACCCCTCAGCGGCTCTT	904
QY	781	GGTCCAGTCTGCTCTCTCC--TCAAAGGACGACGAGCTTGTTCATGATTTTCCATTC-	837
Db	905	GCTCCAGACCTGCCCCCTCTCTTAAGAGGCTGCTGGGCTGTTCACGTGTTTCCATCCC	964
QY	838	-----AAGAAGCTATCTTGTCTCTTAACATCCATCCACACATATCACCCTC	891
Db	965	ACATAAATACAGTATTTCCCACTCTTATCTTACAACTCCCCACCGCCACTCTCCACCTC	1024
QY	892	ACTAGCTCCCAAGCCCCCTAC-----TTATCCCTATCTCCCCACCCACT	936
Db	1025	ACTAGCTCCCAAGCCCCCTAC-----TTATCCCTATCTCCCCACCCACT	1084
QY	937	CACCGACACACGCTTTATGACTTTGTGCAC-----	968
Db	1085	CAGACCCCCAGGCTCTTGTCTTACCTGTACTCTGTGGGCAAGATGGGTCCAGAGACCC	1144
QY	969	-----CAGGCACTGATGATGGGCTGGACCTGGTGGGAGGAGCAGACAGCTGGGACTAG	1023
Db	1145	CACCTTCAGGCACTAAGAGGGGCTGGACCTGGGAGGAGCAGACAGCTGGGACTAG	1204
QY	1024	GCCAGAACTTCCCAACTGTGAGGGGAGAGAGCTGGGAGCAAGCTCTCTCTGGA----	1079
Db	1205	GCCAGGAGTCTCCCAATGTGAGGGGCGAGA-AACAAGCAAGCTCTCTCTGGAATTC	1263
QY	1080	CCTGTGGAATTTTGA--AGATCTATTTTATTTATTTATGTGACAAATGT--TAAT	1134
Db	1264	CCTGTGGAATTTTGA--AGATCTATTTTATTTATTTATGTGACAAATGTGATTAAT	1323
QY	1135	GGATTTTAAGAGATAATATCATGA	1159
Db	1324	GGATTTTAATGATAATAGTCATAA	1348
RESULT 5			
LOCUS	AF055872	1368 bp	linear PRI 04-MAY-1998
DEFINITION	Homo sapiens Apo3/DR3 ligand (Apo3L) mRNA, complete cds.		
ACCESSION	AF055872		
VERSION	AF055872.1	GI:3108230	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens.		
REFERENCE	Enkayrta; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 1368)		
AUTHORS	Marsters,S.A., Sheridan,J.P., Pitti,R.M., Brush,J., Goddard,A. and Ashkenazi,A.		
TITLE	Identification of a ligand for the death-domain-containing receptor Apo3		
JOURNAL	Curr. Biol. 8 (9), 525-528 (1998)		
MEDLINE	9828355		
PUBMED	9503043		
REFERENCE	2 (bases 1 to 1368)		
AUTHORS	Marsters,S.A., Sheridan,J.P., Pitti,R.M., Brush,J., Goddard,A. and Ashkenazi,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-MAR-1998) Molecular Oncology, Genentech, 1 DNA Way, South San Francisco, CA 94080, USA		
FEATURES	Location/Qualifiers		

	source	1..1368
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/chromosome="17"	
	/map="17p13"	
	1..1368	
	/gene="APO3L"	
	58..807	
	/function="binds to the death-domain-containing receptor APO3/DR3"	
	/note="TNF homolog"	
	/codon_start=1	
	/product="APO3/DR3 ligand"	
	/protein_id="AAC39724.1"	
	/db_xref="GI:3108231"	
	/translation="MAARSSQRRRGREGPEPTALLPLATGLALACTLLAVSISGSRSLASQAEPAOEIVAEEDODPSLEINPDEESODAPFLNRYVRRSAPKGRTRARRAIAAHYEVHPRPDGNAGVDVGYSGEERINSSELPRLNRIOGETIYRACIIYLICQVHFDEGKAIVYLKLIDLVDGVLAIRCFEEFSATPAASSLPOLNLCOVSLDLRPSSLRIRLTPMHLNAPELYFGLEYFH"	
	RPGSLRIRLTPMHLNAPELYFGLEYFH"	
	BASE COUNT 272 a 443 c 389 g 264 t	
	ORIGIN	
	Query Match 53.8%; Score 628.6; DB 9; Length 1368;	
	Best Local Similarity 76.2%; Pred. No. 7.5e-155;	
	Matches 933; Conservative 0; Mismatches 219; Indels 73; Gaps 10;	
Oy	2 GGCCTGAGCCTGGCGCTGGCGCTGGCCCTTGCCCTTCCTGTGCTGGTCAAGCCTG 61	
Dd	130 GGCGTGGGCTTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGG 189	
Oy	62 GGGAGCTGGGCAACGCTGTCTGGCCAGAGAGCCTTCTCAGAGAGCTGACAGAGAGAC 121	
Dd	190 GGGAGCCGGGCACTGCTGTCCGCCAGAGAGCCTGCCAGAGAGAGCTGGTGGCAGAGAG 249	
Oy	122 CGCGGGAGACCCTCGAACTGAATCCCCAGACAGAGAGAAAGCCAGAGATGTGTACTTTC 181	
Dd	250 GACCAGAGCCGCTCGGAAGTGAATCCCAGACAGAGAGAAAGCCAGAGATCTTCCGCTTTTC 309	
Oy	182 TTGGAACAATACTAGTCCGGCCTCGAAGAATGCTCTTAAGGCGGGAAGCGCGGCTCGC 241	
Dd	310 CTGAACCGACTTAGTTCGGCCTCGCAGAAAGTCACTTAAGGCGGGAAGAAACAGGCGCTCA 369	
Oy	242 CGAGCTATTGCAAGCCCATTTATGAGGTTCACTCTGGCCAGAGACAGATGGAGCACAGCA 301	
Dd	370 AGAGGAGTCCGAGCCCATTTATGAGTTCATCCAGACACCTGACAGAGAGAGAGCAGGCA 429	
Oy	302 GGTGTGATGGGACAGTGAAGTGGCTGGGAGAGAGACCAAAATACAAGCTCCAGCCCTCTG 361	
Dd	430 GGTGTGAGAGGGGACAGTGAAGTGGCTGGGAGAGAGACCAAAATACAAGCTCCAGCCCTCTG 489	
Oy	362 GCCTCGAAGCCGACATTTAGGGGAATTTACAGCATACAGAGGCTGGGCTCTACTACCTGTAC 421	
Dd	490 GCCTCAACCCGCAATGGGGAGTTTAAAGTCACCCGGGCTGGGCTCTACTACCTGTAC 549	
Oy	422 TGTCAAGTCACTTGTATGATGAGGAAGAGCGTGTCACTCAAGCTGAAGTGTGTGTAAC 481	
Dd	550 TGTCAAGTCACTTGTATGATGAGGAAGAGCGTGTCACTCAAGCTGAAGTGTGTGTAAC 609	
Oy	482 GGTGTGAGGCGGCTGGCTGGAGAGAAATTCAGCCACAGCAGCAAGCTCTCTGAGG 541	
Dd	610 GGTGTGAGGCGGCTGGCTGGAGAGAAATTCAGCCACAGCAGGAGGAGGAGTTCCTCGGG 669	
Oy	542 CCCCAGCTCCGTTTGTGCCAGGTGTCTGGGCTGTGGCGCTCGGCGCAAGGCTCTCCCTT 601	
Dd	670 CCCCAGCTCCGCTTGGCCAGGTTGTGGGCTGTGGCGCTCGGCGCAAGGCTCTCCCTT 729	
Oy	602 CGGATCCGACCCCTCCGCGGCTCATCTTAAGGCTGCCGCCCTTCTCAACTACTTGTGA 661	
Dd	730 CGGATCCGACCCCTCCGCGGCTCATCTTAAGGCTGCCGCCCTTCTCAACTACTTGTGA 789	
Oy	662 CTCCTTCAAGTTCACTGAAGGGGCTTGGCTCTCCAGATTCCTTAAACTTCCCTGGCTCC 721	

```

|||||
Db 790 CTTCTTCACGTTCACTGAGGGCCCTGGTCTCCCGCACTGCTCCAGGCTGCCGCTTC 849
Oy 722 AGGACATACACACACTCCCTACCCACACCCCACTCCCTCCACCCCTC-GCTGCTCTT 780
Db 850 -----CTTCGACAGCTCTCTGGGCAACCGGTCCTCCCTCCGACCCCTCAGCCGCTCTT 904
Oy 781 GGTCCAGTCTCTCTCTCC--TCAAAAGCAGCCAGCAAGCTTGTTCACATGTTTCCATTCG- 837
Db 905 GCTCCAGACCTGCCCCCTCTCTAGAGGCTGCTGGGCTGTTCACAGTGTTCATCCATCC 964
Oy 838 -----ACAGACGTATCCCTTCTTCTTAACATCCCATCCACACCAACTATCCACCTC 891
Db 965 ACATAAATATACATATATCCACCTTATCTTACACTCCCCACGCGCCACTCTCCACCTC 1024
Oy 892 ACTAGCTCCCAAGCCCTAC-----TTATCCCTGACTCCCCACCCACT 936
Db 1025 ACTAGCTCCCAAGCCCTGAGGCCCCCAGATGATGATGATGATGATGATGATGATGATG 1084
Oy 937 CACCGGACCACTGTTTATGACTTGTGAC----- 968
Db 1085 CAGACCCCGAGTCAATTTGTTGACTGCTGCTGAGGCAAGATGGGTCCAGAAAGCC 1144
Oy 969 -----CAGGCACTGATGAGGCTGAGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1023
Db 1145 CACTTCAGGCACTAGAGGGGCTGACCTGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1204
Oy 1024 GCCAGAGTCCCAACTGTGAGGGGAGAGAGCTGCTCCCTGGA-----TC 1079
Db 1205 GCCAGAGTCCCAACTGTGAGGGGAGAG--AACAAAGCAAGCTCTCCCTGGAAGTTC 1263
Oy 1080 CCGTGGATTTTGA--AGATACATTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1134
Db 1284 CCGTGGATTTTGA--AGATACATTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 1323
Oy 1135 GGATATTAAAGAGATATATCATGA 1159
Db 1324 GGATATTAAAGATATATCATGA 1348

RESULT 6
AF030099 1306 bp mRNA Linear PRI 20-DEC-1997
LOCUS Homo sapiens TWEAK mRNA, complete cds.
DEFINITION AF030099
ACCESSION AF030099
VERSION AF030099.1 GI:2707218
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1306)
AUTHORS Chicheportiche,Y., Bourdon,P.R., Xu,H., Hau,Y.M., Scott,H.,
Hession,C., Garcia,I. and Browning,J.L.
TWEAK, a new secreted ligand in the tumor necrosis factor family
that weakly induces apoptosis
J. Biol. Chem. 272 (51), 32401-32410 (1997)
JOURNAL MEDLINE
PUBMED 9405449
REFERENCE 2 (bases 1 to 1306)
AUTHORS Bourdon,P., Hession,C., Tizard,R. and Browning,J.
TITLE Direct Submission
JOURNAL Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,
Cambridge, MA 02142, USA
FEATURES
Location/Qualifiers
source 1..1306
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
/issue_type="tonsil"
18..767
/note="Ligand in the TNF family; secreted protein; start

```

```

codon not verified experimentally"
/codon_start=1
/product="TWEAK"
/protein_id="AAC51923.1"
/db_xref="GI:2707219"
/translation="MAARSRORRRRGEDEPTALLVPLALGLIALCLALLAVSL
GSRAISLAQEPQAEELVAEEDQPSLEINPQEEQDAPFLNRLVPRRSAPKGRKR
ARRAIAHVEVPRPDGADQAGVDEARINSSEPLRNROIGEFIVRAGL
YLYCOVHFDEGRKAVYLLKDLVDGVATLRLEEFESATVAASISLQPOLCOVSSLAL
RPSLSLRIRLPMHAKAPFLYFEGFYOH"
BASE COUNT 247 a 434 c 368 g 257 t
ORIGIN

```

```

Query Match 53.4%; Score 624; Db 9; Length 1306;
Best Local Similarity 76.0%; Pred. No. 1,2e-153;
Matches 929; Conservative 0; Mismatches 220; Indels 73; Gaps 10;

```

```

Oy 2 GTGCTGAGCTTGGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 61
Db 90 GCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 149
Oy 62 GGGAGCTGGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 121
Db 150 GGGAGCTGGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 209
Oy 122 GCGCGGAGCCCTCGAATGATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 181
Db 210 GACCAAGACCGGTGGAAGTGAATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 269
Oy 182 TTGGAACAATAGTCCGCTCGAAGAGTCTCTAAAGGCGGAGAGAGAGAGAGAGAGAGAGAGAG 241
Db 270 CTGAACCGAGTATGTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 329
Oy 242 CGAGCTATGGAGCCATATATGAGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301
Db 330 AGAGGATCGAGAGCCATATATGAGTTATCTACAGAGCTGAGAGAGAGAGAGAGAGAGAGAGAG 389
Oy 302 GGTGTGATGGAGAGTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
Db 390 GGTGTGAGAGAGAGTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 449
Oy 362 CGCTACGAGCGCGATTTGGGGAATTTACATCATAGAGGCTGGGCTCTACTACTCTGATC 421
Db 450 CGCTACAGCGCGATTTGGGGAATTTATCTACCGCGGCTGGGCTCTACTACTCTGATC 509
Oy 422 TGTCAAGTGCATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 481
Db 510 TGTCAAGTGCATTTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 569
Oy 482 GGTGTGCTGGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 541
Db 570 GGTGTGCTGGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 629
Oy 542 CCCCAGCGCTGTTGGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 601
Db 630 CCCCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 689
Oy 602 CGGATCGGACCTCCCTGCGCTCATTTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 661
Db 690 CGGATCGGACCTCCCTGCGCTCATTTAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 749
Oy 662 CTCTTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 721
Db 750 CTCTTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 809
Oy 722 AGGAGATACACACACTCCCTACCCACACCCCACTCCCTCCACCCCTC-GCTGCTCTT 780
Db 810 -----CTTCGACAGCTTCTGGGACCGGTCCTCTTGGCCCACTCCAGCGGCTCTT 864
Oy 781 GGTCCAGTCTCTCTCTCC--TCAAAAGCAGCCAGCAAGCTTGTTCACATGTTTCCATTCG- 837
Db 865 GGTCCAGACCTGCCCCCTCTCTAGAGGCTGCTGGGCTGTTCACAGTGTTCATCCATCC 924

```

[illegible]

Eukaryota; Metazoa; Cnidaria; Cnidaria; Vertebrata; Euteleostomi;

REFERENCE  
AUTHORS

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

1 (bases 1 to 138792)

Murphy D.M., Adams C., Adio-Oduola, B., Ali-Osman, F.R., Allen C.,  
Albrooks, S.L., Amarantunga, H.C., Are, J.R., Ayele, M., Banks, T.,  
Bardar, J., Benton, J., Blinag, K., Blankenburg, K., Bonni, D.,  
Bouck, J., Bowler, S., Burke, M., Brown, M., Bryant, N.P.,  
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,  
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,  
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
DeLaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,  
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,  
Homel, F., Howard, S., Huber, J., Huiyk, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karlsom, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kratovic, U., Kuresh, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H.,  
Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mel, G., Melker, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, N.,  
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogum, M., Okunolu, G.,  
Orangunye, N., Oviado, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G.,  
Scherer, S., Scott, G., Shen, H., Shooshari, N., Sisson, I.,  
Sodergren, E., Sotake, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telitod, B., Thomas, N., Thomas, S.,  
Usmani, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorilla, S., Nelson, D.,  
Weinstock, G. and Gibbs, R.

TITLE  
JOURNAL

Direct Submission

2 (bases 1 to 138792)

Worley, K.C.

Direct Submission

Submitted (25-Apr-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 138792)

Worley, K.C.

Direct Submission

Submitted (18-Jul-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA

On Jul 14, 2002 this sequence version replaced gi.20303440.

Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)

Project Information

Center project name: GUXG

Center clone name: CH230-320N23

Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap: version 0.990329

Consensus quality: 112720 bases at least Q40

Consensus quality: 116666 bases at least Q30

Consensus quality: 119165 bases at least Q20

Consensus quality: 119165 bases at least Q20

Consensus quality: 119165 bases at least Q20

Consensus quality: 119165 bases at least Q20

Consensus quality: 119165 bases at least Q20

Consensus quality: 119165 bases at least Q20

Consensus quality: 119165 bases at least Q20

Consensus quality: 119165 bases at least Q20

Consensus quality: 119165 bases at least Q20

\* NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 32 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1	1022:	contig of 1022 bp in length
1023	1122:	gap of unknown length
1123	2394:	contig of 1272 bp in length
2395	2494:	gap of unknown length
2495	3663:	contig of 1169 bp in length
3664	3763:	gap of unknown length
3764	4879:	contig of 1116 bp in length
4880	4979:	gap of unknown length
4980	6787:	contig of 1808 bp in length
6788	6887:	gap of unknown length
6888	8192:	contig of 1305 bp in length
8193	8292:	gap of unknown length
8293	9461:	contig of 1169 bp in length
9462	9561:	gap of unknown length
9562	11435:	contig of 1874 bp in length
11436	11535:	gap of unknown length
11536	13535:	contig of 2000 bp in length
13536	13635:	gap of unknown length
13636	17323:	contig of 3688 bp in length
17324	17423:	gap of unknown length
17424	20406:	contig of 2983 bp in length
20407	20506:	gap of unknown length
20507	23445:	contig of 2939 bp in length
23446	23545:	gap of unknown length
23546	25600:	contig of 2055 bp in length
25601	25700:	gap of unknown length
25701	28049:	contig of 2349 bp in length
28050	28149:	gap of unknown length
28150	30474:	contig of 2325 bp in length
30475	30574:	gap of unknown length
30575	35498:	contig of 424 bp in length
35499	35598:	gap of unknown length
35599	39284:	contig of 3686 bp in length
39285	39384:	gap of unknown length
39385	41970:	contig of 2586 bp in length
41971	42070:	gap of unknown length
42071	45659:	contig of 3589 bp in length
45660	45759:	gap of unknown length
45760	50332:	contig of 4573 bp in length
50333	50432:	gap of unknown length
50433	54695:	contig of 4263 bp in length
54696	54795:	gap of unknown length
54796	58289:	contig of 3494 bp in length
58290	58389:	gap of unknown length
58390	63599:	contig of 5210 bp in length
63600	63699:	gap of unknown length
63700	70305:	contig of 6606 bp in length
70306	70405:	gap of unknown length
70406	76123:	contig of 5718 bp in length
76124	76223:	gap of unknown length
76224	84961:	contig of 8738 bp in length
84962	85061:	gap of unknown length
85062	93614:	contig of 8553 bp in length
93615	93714:	gap of unknown length
93715	103352:	contig of 9638 bp in length
103353	103452:	gap of unknown length
103453	110299:	contig of 6647 bp in length
110300	110399:	gap of unknown length
110400	116180:	contig of 5781 bp in length
116181	116280:	gap of unknown length
116281	125484:	contig of 9204 bp in length
125485	125584:	gap of unknown length









REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.  
1 (bases 1 to 215795)  
Akter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P., Lee, L.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B., Marquies, E.H., Mastello, C., Maskeri, B., Mastrian, S.D., McLoskey, J.C., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Schueler, M.G., Statitipop, S., Thomas, J.W., Thomas, P., Touchman, J.W., Tsurgou, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 215795)  
Green, E.D.

TITLE  
JOURNAL  
AUTHORS

Direct Submission  
Submitted (17-JUL-2002) NIH Intramural Sequencing Center, 8717  
Groveont Circle, Gaithersburg, MD 20877, USA

## COMMENT

Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: <http://www.nisc.nih.gov>  
Contact: nisc.zoehngf1.nih.gov  
Project Information  
Center project name: cms  
Center clone name: 145D13

Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; Version 0.990119  
Consensus quality: 203929 bases at least Q40  
Consensus quality: 209181 bases at least Q30  
Insert size: 190000; agarose-1p  
Insert size: 214395; sum-of-ctrls  
Quality coverage: 7.30x in Q20 bases; agarose-1p  
Quality coverage: 6.47x in Q20 bases; sum-of-ctrls

NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 3593: contig of 3593 bp in length  
\* 3594 3693: gap of unknown length  
\* 3694 8001: contig of 4308 bp in length  
\* 8002 8101: gap of unknown length  
\* 8102 15639: contig of 7538 bp in length  
\* 15640 15739: gap of unknown length  
\* 15740 24775: contig of 9036 bp in length  
\* 24776 24875: gap of unknown length  
\* 24876 30150: contig of 5275 bp in length  
\* 30151 30250: gap of unknown length  
\* 30251 39212: contig of 8962 bp in length  
\* 39213 39312: gap of unknown length  
\* 39313 49342: contig of 10030 bp in length  
\* 49343 49442: gap of unknown length  
\* 49443 60956: contig of 11514 bp in length  
\* 60957 61056: gap of unknown length  
\* 61057 74520: contig of 13464 bp in length  
\* 74521 74620: gap of unknown length  
\* 74621 89522: contig of 14902 bp in length  
\* 89523 89622: gap of unknown length  
\* 89623 105537: contig of 15915 bp in length  
\* 105538 105637: gap of unknown length  
\* 105638 121197: contig of 15560 bp in length  
\* 121198 121297: gap of unknown length  
\* 121298 138883: contig of 17586 bp in length

FEATURES  
source  
\* 138884 138983: gap of unknown length  
\* 138984 166802: contig of 27819 bp in length  
\* 166803 166902: gap of unknown length  
\* 166903 215795: contig of 48693 bp in length.  
Location/Qualifiers  
1. 215795  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="RP43-145D13"  
/clone\_11b="RP43"  
1. 3593  
/note="assembly\_fragment"  
3694. 8001  
/note="assembly\_fragment"  
8102. 15639  
/note="assembly\_fragment"  
15740. 24775  
/note="assembly\_fragment"  
24876. 30150  
/note="assembly\_fragment"  
30251. 39212  
/note="assembly\_fragment"  
39313. 49342  
/note="assembly\_fragment"  
49443. 60956  
/note="assembly\_fragment"  
61057. 74520  
/note="assembly\_fragment"  
74621. 89522  
/note="assembly\_fragment"  
89623. 105537  
/note="assembly\_fragment"  
105638. 121197  
/note="assembly\_fragment"  
121298. 138883  
/note="assembly\_fragment"  
138984. 166802  
/note="assembly\_fragment"  
166903. 215795  
/note="assembly\_fragment"

BASE COUNT 52160 a 54859 c 53752 g 53513 t 1511 others  
ORIGIN  
Query Match 26.0%; Score 304; DB 2; Length 215795;  
Best Local Similarity 71.7%; Pred. No. 7.9e-69;  
Matches 581; Conservative 0; Mismatches 155; Indels 74; Gaps 11;  
QY 425 CAGGTGACCTTGTGATGAGGAAAGCTGTCTACTGTAAGCTGAGCTGTGTAACGCT 484  
|||||  
DB 199456 CAGGTGACCTTGTGATGAGGAAAGCTGTCTACTGTAAGCTGAGCTGTGTAACGCT 199515  
QY 485 GTGTGAGCCCTGCGTCTGTGTAAGATTTCTAGACACAGCAAGCAAGCTCTCTGGGCC 544  
|||||  
DB 199516 GTGTGAGCCCTGCGTCTGTGTAAGATTTCTAGACACAGCAAGCTCTCTGGGCC 199575  
QY 545 CAGCTCCGTTTGTGACAGGTGTCTGAGCTGTGCGCTGCGGACAGAGGCTCTCCCTGG 604  
|||||  
DB 199576 CAGCTCCGTTTGTGACAGGTGTCTGAGCTGTGCGCTGCGGACAGAGGCTCTCCCTGG 199635  
QY 605 ATTCGACCCCTCCCTGAGCTCATTTAAGGCTGCGCCCTTCTCTAAGCTTCTTGGACTC 664  
|||||  
DB 199636 ATTCGACCCCTCCCTGAGCTCATTTAAGGCTGCGCCCTTCTCTAAGCTTCTTGGACTC 199695  
QY 665 TTTCAGATTCATGAGAGGCTGTCTCTCCAGATTTCTTAACATTTCTCCCTGGCTCCAG 724  
|||||  
DB 199696 TTTCAGATTCATGAGAGGCTGTCTCTCCAGATTTCTTAACATTTCTCCCTGGCTCCAG 199752  
QY 725 AGCATACACACCTTCACACCCACCCACCTCCGACCCCTC-GCTGCTGCTGTGT 783  
|||||  
DB 199753 --CTTCGACAGCTCTCTGTGGGCAACCGGCTCTCTGCCCCACCTTCAGCTGCTTTTGT 199810  
QY 784 CCAGTCTGTCTCTCC--TCAAAGGACGACAGCTTTTTCACAG--TTTTCATTCCACA 840  
|||||

Db 199811 CCAGACCTGCCCTCCCTCTAGAGGCTGCCCTGTCACGTTTTCATCCACCA 199870  
 QY 841 GACGTA-----TCCTTGTCTCTTAAATCCATCCACCAACTATCCAGCTAC 893  
 Db 199811 TAAATACAGTATCCCACTCTATCTTACAACTCCCCACGCCCACTCTCCACCTAC 199930  
 QY 894 TAGTCCCAAGCCCTAC-----TTATCCCTGACTCCCACTCCACTCA 938  
 Db 199931 TAGTCCCAAGCCCTACCTTGTAGAGCCCAAGTATCTGACTCCCTCCGACCA 199990  
 QY 939 CCGACCACTGTTTATGACTTGTGAC----- 968  
 Db 199991 GACCCGAGGCTGTTTCTACTCTACTCTGAGGCAAGATGGTCCAGAACCCCA 200050  
 QY 969 ---GAGCAGTACAGTGGGCTGAGCTGTGTGAGCAAGAACCTGGAGTACGC 1025  
 Db 200051 CTTAGGACACTTACAGGGGCTGAGCTGTGTGAGCAAGAACCTGGAGTACGC 200110  
 QY 1026 CAGAGTTCCTCACTGTGAGGGGAGAGACGTCGCTCCCTGGA---TCCC 1081  
 Db 200111 CAGAGTTCCTCACTGTGAGGGGAGAGACGTCGCTCCCTGGA---TCCC 200169  
 QY 1082 TGTGATTTTGA---AGATCTATTTTATTTATTTATTTGACAAATGT---TAAATGG 1136  
 Db 200170 TGTGATTTTGA---AGATCTATTTTATTTATTTATTTGACAAATGT---TAAATGG 200229  
 QY 1137 ATATTAAAGATTAATCATGATTTCTCT 1166  
 Db 200230 ATATTAAAGATTAATCATGATTTCTCT 200259

RESULT 12  
 AC016876/c 177703 bp DNA linear HTG 06-AUG-2002  
 LOCUS Homo sapiens clone Rpl1-186B7, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 10  
 DEFINITION unordered pieces.  
 ACCESSION AC016876  
 VERSION AC016876.5 GI:21313830  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 177703)  
 AUTHORS Birren,B., Nussbaum,C. and Lander,E.  
 TITLE Homo sapiens chromosome, clone Rpl1-186B7  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 177703)  
 AUTHORS Birren,B., Linton,L., Nussbaum,C., Lander,E., Allen,N., Anderson,M.,  
 Baldwin,J., Batra,N., Beckert,R., Boguslavsky,L., Boukhgalter,B.,  
 Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,  
 Cooke,P., Dearellano,K., Dewar,K., Domiano,M., Donelan,L., Doyle,M.,  
 Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,  
 Galagan,J., Gardyna,S., Grant,G., Hages,B., Heathford,A., Horton,L.,  
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
 Lehotzky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,  
 McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,  
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
 Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Tsafaye,S., Tirelli,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
 Wyman,D., Ye,W.-J., Zimmer,A. and Zody,M.

TITLE Direct Submission  
 JOURNAL Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 REFERENCE 3 (bases 1 to 177703)  
 AUTHORS Birren,B., Nussbaum,C., Lander,E., Allen,N., Anderson,M.,  
 Batra,N., Beckert,R., Boguslavsky,L., Boukhgalter,B.,  
 Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,  
 Cooke,P., Dearellano,K., Dewar,K., Domiano,M., Donelan,L., Doyle,M.,  
 Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J.,  
 Gardyna,S., Grant,G., Hages,B., Heathford,A., Horton,L.,  
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
 Lehotzky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,  
 McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,  
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
 Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Tsafaye,S., Tirelli,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
 Wyman,D., Ye,W.-J., Zimmer,A. and Zody,M.

TITLE Direct Submission  
 JOURNAL Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 REFERENCE 1 (bases 1 to 177703)  
 AUTHORS Birren,B., Linton,L., Nussbaum,C., Lander,E., Allen,N., Anderson,M.,  
 Baldwin,J., Batra,N., Beckert,R., Boguslavsky,L., Boukhgalter,B.,  
 Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,  
 Cooke,P., Dearellano,K., Dewar,K., Domiano,M., Donelan,L., Doyle,M.,  
 Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J.,  
 Gardyna,S., Grant,G., Hages,B., Heathford,A., Horton,L.,  
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
 Lehotzky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,  
 McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,  
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
 Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Tsafaye,S., Tirelli,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
 Wyman,D., Ye,W.-J., Zimmer,A. and Zody,M.

TITLE Direct Submission  
 JOURNAL Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 COMMENT On Jun 2, 2002 this sequence version replaced g1:15421989.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WtBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L3849  
 Center Clone name: 186\_B\_7

NOTE: This is a 'working draft' sequence. It currently  
 consists of 10 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1 48645: contig of 48645 bp in length  
 \* 48646 48745: gap of 100 bp  
 \* 48746 58277: contig of 9532 bp in length  
 \* 58278 58377: gap of 100 bp  
 \* 58378 65804: contig of 7427 bp in length  
 \* 65805 65904: gap of 100 bp  
 \* 65905 79793: contig of 13889 bp in length  
 \* 79794 79893: gap of 100 bp  
 \* 79894 99493: contig of 19600 bp in length  
 \* 99494 99593: gap of 100 bp  
 \* 99594 111049: contig of 11456 bp in length  
 \* 111050 111149: gap of 100 bp  
 \* 111150 125020: contig of 13871 bp in length  
 \* 125021 125120: gap of 100 bp  
 \* 125121 145109: contig of 19969 bp in length  
 \* 145110 145209: gap of 100 bp  
 \* 145210 169458: contig of 24249 bp in length  
 \* 169459 169558: gap of 100 bp  
 \* 169559 177703: contig of 8145 bp in length.

FEATURES  
 source  
 1. 177703  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="Rpl1-186B7"  
 /clone\_1b="RPl-11 Human Male BAC"  
 BASE COUNT 42862 a 45967 c 46568 g 41254 t 1052 others  
 ORIGIN

Query Match 26.0% Score 303.8; DB 2; Length 177703;  
 Best Local Similarity 71.0% Pred. No. 8.7e-69;  
 Matches 574; Conservative 0; Mismatches 162; Indels 73; Gaps 10;

QY 425 CAGGTGACCTTGTATGAGGAGGAGGCTGTCTACTGAGAGCTGTGCTGAGACGT 484  
 Db 37314 CAGGTGACCTTGTATGAGGAGGAGGAGGCTGTCTACTGAGAGCTGTGCTGAGACGT 37255  
 QY 485 GTGCTGCCCTCGGCTCTCGGAGAGATTTCTACGCCACAGCAGACGCTCTCGGCGCC 544









GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 8, 2003, 02:27:27 ; Search time 1333.52 Seconds

(without alignments)

3449.164 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 1444

Sequence: 1 MSLLDFEISARRLEPRSLG.....PMAHKAPFLTYGILQVH 284

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL-frame+g2n.model -DEV-xlp  
-O=/cgn2.1/USPTO.spool/US09245198/runat\_06052003\_170239\_9628/app\_query.fasta\_1.846  
-DB-EST -QFMT-fastlap -SUFFIX-first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=-1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST=45  
-DOCLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTMT=pic -NORMext -HEAPSIZ=500 -MTLEN=0 -MAXLEN=200000000  
-USER=US09245198.ecgn.1.1.235.6runat.06052003\_170239\_9628 -NCPV=6 -ICPV=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_hum:\*  
20: em\_gss\_hum:\*  
21: em\_gss\_pln:\*  
22: em\_gss\_vrt:\*  
23: em\_gss\_fun:\*  
24: em\_gss\_mam:\*  
25: em\_gss\_mus:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1087.5	73.3	731	13	B1871711
2	1060.5	73.4	728	13	B1870393
3	979	67.8	777	13	B1819200
4	962	66.6	834	13	B1766766
5	934	64.7	828	13	B1596681
6	913	63.2	918	12	BF577781
7	909.5	63.0	948	14	BQ707185
8	904	62.6	940	14	BQ884231
9	825	57.1	561	10	AW763237
10	809	56.0	963	14	BQ671259
11	765	53.0	1071	14	BQ921213
12	753	52.1	951	14	BQ674188
13	742.5	51.4	785	13	B1762908
14	738	51.1	587	13	BG686319
15	707	49.0	1033	11	AK020909
16	652	45.2	584	10	AW917574
17	648.5	44.9	698	13	B1906850
18	647	44.8	894	13	B1908274
19	645.5	44.7	474	13	B1965174
20	604	41.8	445	9	AA870722
21	598.5	41.4	471	9	AA221610
22	582.5	40.3	542	12	BF041509
23	582	40.3	650	12	BG404836
24	572.5	39.6	404	12	BF044430
25	562	38.9	360	10	BE654876
26	520.5	36.0	493	10	BE307031
27	506	33.0	538	12	BF821434
28	499	34.6	558	13	BW484863
29	493	34.1	298	14	BM688946
30	476	33.0	441	13	B1967060
31	466	33.3	315	12	BF466521
32	441	30.5	465	9	A1091441
33	440	30.5	581	13	B1738634
34	421	29.2	345	14	R55379
35	420	29.1	278	10	AA417023
36	403	27.9	1187	14	BQ053284
37	395.5	27.4	571	12	BF073881
38	391	27.1	639	12	BB642326
39	388.5	26.9	1926	14	BM906056
40	363	25.1	400	9	A1152313
41	362	25.1	342	9	AA637970
42	353	24.4	910	12	BG110063
43	352	24.4	456	13	B1966255
44	347	24.0	487	10	AW320117
45	320	22.2	533	10	BE628951

#### ALIGNMENTS

RESULT 1  
LOCUS B1871711  
DEFINITION 603395825F1 NIH\_MGC\_90 Homo sapiens CDNA clone IMAGE:5405478 5',  
ACCESSION B1871711  
VERSION B1871711.1 GI:16045386  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 731)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)





```

OY 104 SerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeu 123
    |||||||
DB 122 TCGGACTGATCCCGCAGACAGAAAGAACCCAGATCCGCCCTTCTCTGAACGACATA 181
OY 124 ValArgProArgArgSerAlaProLysGlyArgLysThrArgAlaArgAlaLeala 143
    |||||||
DB 152 GTTGCGCTCCGAGAAAGTGCACCTAAAGCCGGAACACGGGGCTCGAAGACGATCGCA 241
OY 144 AlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGly 163
    |||||||
DB 242 GCCCATTTATGAAATTCATCCACGACCTGACAGAGAGAGCCAGGAGTGTGGACGG 301
OY 164 ThrValSerGlyTyrGluGluAlaArgLysLeuAsnSerSerProLeuArgTyrAsnArg 183
    |||||||
DB 302 ACAGTACGATGCTGGAGAGAACCCAGAAATCAACAGCTCCAGCTCTGCTACAAACCGC 361
OY 184 GlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHis 203
    |||||||
DB 362 CAGATCGGGAGTTTATAGTACACCGGGCTGGCTCTACTACTGACTGACGAGTGCAC 421
OY 204 PheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGlyValLeuAla 223
    |||||||
DB 422 TTTGATGAGGGAGAGCTGTCTACTGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTG 481
OY 224 LeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArg 243
    |||||||
DB 482 CTGGCTGCTGCTGGAGAAATTCACGACCTGCGGCGCAAGTCTCCCTGGGGCCCAAGCTCCG 541
OY 244 LeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThr 263
    |||||||
DB 542 CTCTCCAGAGTGTCTGGGCTGTGGCCCTGCGGCGCAGAGTCC-TCCCTGGAGATCCGACAC 600
OY 264 LeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 283
    |||||||
DB 601 CTCCCTGGGGCCATCTCAAGGCTGC-CCCTTCTCCTACCTACTTGGAGCTTCCAGGT 659
OY 284 His 284
    |||
DB 660 CAC 662

RESULT 3
LOCUS B1819200 777 bp mRNA linear EST 04-OCT-2001
DEFINITION 603034614P1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175698 5',
    mRNA sequence.
ACCESSION B1819200
VERSION B1819200.1 GI:15930750
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 777)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
    Email: cgabs-remail.nih.gov
    Tissue Procurement: Life Technologies, Inc.
    cDNA Library Preparation: Life Technologies, Inc.
    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
    DNA Sequencing by: Incyte Genomics, Inc.
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/LNLN at:
    http://image.llnl.gov
    Plate: L14M11437 row: 1 column: 03
    High quality sequence stop: 759.
FEATURES
    Source
        1..777
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:5175698"

```

```

/clone.lib="NIH_MGC_115"
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
PCMV-SpOBS; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH-MGC library."

BASE COUNT 140 a 268 c 215 g 154 t
ORIGIN

Alignment Scores:
Pred. No.: 5,336-83 Length: 777
Score: 979.00 Matches: 188
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 67.80% Indels: 0
DB: 13 Gaps: 0

US-09-245-198A-4 (1-284) x B1819200 (1-777)
OY 97 AlaGluGluAspGlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspPro 116
    |||||||
DB 14 GCAGAGAGAGACAGAGACCCGTCGGAATGAAATCCCAACAGAAAGCCAGAGTCTCT 73
OY 117 AlaProPheLeuAsnArgLeuValArgProArgSerAlaProLysGlyArgLysThr 136
    |||||||
DB 74 GGGCTTCTCTGAACCGACTAGTGTGGCTCCAGAAAGTGCACCTTAAGCCGGAAAAACA 133
OY 137 ArgAlaArgArgAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGly 156
    |||||||
DB 134 CGGGCTCGAAGAGCATCCACCCCATTTATGAAATTCATCCAGACCTGAGAGCGCA 193
OY 157 AlaGlnAlaGlyValAspGlyThrValSerGlyTyrGluGluAlaArgLysAsnSerSer 176
    |||||||
DB 194 GCGACAGCAGGTGTGACGCGAGAGTGTGAGTGTGAGAGAAAGCCAGAAATCAACAGCTCC 253
OY 177 SerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyr 196
    |||||||
DB 254 ACCCTCTGCTGCTACACCGCCAGATCGGGAGTTTATATGTCACCGGGCTGGCTTAC 313
OY 197 TyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu 216
    |||||||
DB 314 TACCTGTAAGTCAAGTGCACCTTGTATGAGGAGGAGCTGTCTACTGAAGCTGACCTTG 373
OY 217 LeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAspSer 236
    |||||||
DB 374 CTGGTGGATGATGTCTGCTGGCCCTGCGCTCCGAGAGAAATTCACGCCCTCGCGGAGT 433
OY 237 SerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGly 256
    |||||||
DB 434 TCCCTCGGGGCCACACTCGCCCTCTGCCAGAGTGTCTGGGGCTTGTGGCCCTGGCGCAGG 493
OY 257 SerSerLeuArgIleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThr 276
    |||||||
DB 494 TCCCTCGGGCGATCGGACCTCCCTCGGGGCCATCTCAAGGCTGCCCCCTCTCCAC 553
OY 277 TyrPheGlyLeuPheGlnValHis 284
    |||||||
DB 554 TACTTCGGACTCTTCAGAGTTTAC 577

RESULT 4
LOCUS B1766766 834 bp mRNA linear EST 25-SEP-2001
DEFINITION 603056866P1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206217 5',
    mRNA sequence.
ACCESSION B1766766
VERSION B1766766.1 GI:15758344
KEYWORDS EST.

```

SOURCE	ORGANISM
REFERENCE	human.
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	1 (bases 1 to 834)
COMMENT	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> . National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: gcgabs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> plate: LLM1517 row: C column: 18 High quality sequence stop: 772.

FEATURES

source

Location/Qualifiers

I . 834

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5206217"
/clone_1lb="NIH-MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note
this is a NIH-MGC Library."
154 a 287 c 223 g 170 t

```

Alignment Scores:	
Pred. No.:	2,46e-81
Score:	962.00
Percent Similarity:	98.97%
Best Local Similarity:	98.97%
Query Match:	66.62%
DB:	13
Length:	834
Matches:	192
Conservative:	0
Mismatches:	1
Indels:	2
Gaps:	0

[illegible]

Qy	211	rleuysleuhspleuValAspGlyValleuAlaleuArgCysleuGluInPheSe	231
Db	363	cctGAAGCTGGACCTGGCTGGTAGAGGTGTCTGGCCCTGGCTCCGAGAGAAATCTC	422
Qy	231	rAlaThrAlaIAsSerSerleuGlyProGlnleuArgLeuCysGlnValSerGlyLeuIe	251
Db	423	AGCCACTCGGGCAATTCCTCCCTGGGCCACAGTCCGCTCGCAGAGTGTCTGGGCTGTT	482
Qy	251	uAlaleuArgProGlySerSerleuArgIleArgThrleuProTrpAlaHisleuLysAl	271
Db	483	GGCCCTGGGGCAGAGTCC-TCCCTCGGATCGCACCCTCCCTGGGCCATCTCAAGGC	541
Qy	271	aAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis	284
Db	542	TGCCCTTCTCTCACCTACTTCGGAGCTTTCACAGGTTTCAAC	581

RESULT	5
B1596681.	
LOCUS	
DEFINITION	B1596681 828 bp mRNA linear EST-07-SEP-2001
ACCESSION	60324225.F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285892.5,
VERSION	mRNA sequence.
KEYWORDS	B1596681 B1596681
SOURCE	B1596681.1 GI:15489620
ORGANISM	EST.
REFERENCE	human.
AUTHORS	Homo sapiens
TITLE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
JOURNAL	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
COMMENT	1 (bases 1 to 828)
	NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .
	National Institutes of Health, Mammalian Gene Collection (MGC)
	Unpublished (1999)
	Contact: Robert Strausberg, Ph.D.

Tissue Procurement: Miklos Palkovits, M.D., Ph.D.  
 cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki  
 Toshiyuki and Piero Carninci (RIKEN)  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
<http://image.lnl.gov>  
 Plate: L1AM1722 row: k column: 13  
 High quality sequence stop: 7776.

```

FEATURES
source
Location/Qualifiers
1..828
/orrganism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5285892"
/clone_1id="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescriptpr (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcagg
); Oligo-dT primed using primer 5'-TTTTTTTTTTTTVVN-3',
size-selected for average insert size 2.3 kb and
normalized to ROF 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NIH_MGC library."
BASE COUNT      155 a      278 c      223 g      172 t
ORIGIN
Alignment Scores:
Pred. No.:      1.13e-78      Length:      828
Score:          934.00      Matches:      179
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.44%      Mismatches: 0
Query Match:    64.68%      Indels:      0
DB:             13      Gaps:      0
OS-09-245-198A-4 (1-284) x B1596681 (1-828)

```

QY 105 GluLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgVal 124  
|||||  
Db 44 GAACGATCCCGACAGAGAAAGAACGACGATCTGGCCCTTCTTACACGACTGAT 103  
QY 125 ArgProArgSerAlaProLysGlyArgGlyThrArgAlaArgAlaIleAla 144  
|||||  
Db 104 CGGCTCCAGAGATGACCTAAAGCGGAAACACAGGCTCGAAGAGCATGCGACGC 163  
QY 145 HisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThr 164  
|||||  
Db 164 CATATGATGATTCATCCACGACCTGGACAGAGCGAGCGGCGTGTGACCGGACA 223  
QY 165 ValSerGlyTyrPgluGluAlaAlaArgIleAsnSerSerSerProLeuArgTyrAsnArgGln 184  
|||||  
Db 224 TTGAGTGGCTGGAGGAGAACCGAATCAACAGCTCCACCTCTGCTACACCGCCAG 283  
QY 185 IleGlyIupheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPhe 204  
|||||  
Db 284 ATCGGGAGATTATAGTACACCGGGCTGGCTCTACTACTGTACTGTACAGTGCATTT 343  
QY 205 AspGluGlyValAlaValTyrLeuLysLeuAspLeuValAspGlyValLeuAlaLeu 224  
|||||  
Db 344 GATGAGGGGAGAGCTGTCTACTGAGAGCTGCTGCTGCTGCTGCTGCTGCTG 403  
QY 225 ArgCysLeuGluGluPheSerAlaThrAlaIleSerSerLeuGlyProGlnLeuArgLeu 244  
|||||  
Db 404 CGCTGCCGAGAGATTTCTACGCCCTCGGCTGCTGCTGCTGCTGCTGCTGCTG 463  
QY 245 CysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThrLeu 264  
|||||  
Db 464 TGCCAGGTGTCTGGCTGTGGCTGGCGGAGGCTCTCCCTCGGATCGGACCCCTC 523  
QY 265 ProTrrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 284  
|||||  
Db 524 CCTGGGCCCATCTCAAGGCTGCCCTTCTCTACCTACTGCGACTCTTCAGGTTGAC 583

RESULT 6  
BF577781  
LOCUS 918 bp mRNA linear EST 12-DEC-2000  
DEFINITION M20209.2080F1 NCI\_CGAP\_C024 Mus musculus cDNA clone IMAGE:4206595 5',  
6R20209.2080F1  
ACCESSION BF577781  
VERSION BF577781.1 GI:11651493  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
JOURNAL NIH-MGC http://mgi.nci.nih.gov/  
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: rgs@pds-remail.nih.gov  
Tissue Procurement: Jeffrey E. Green, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMU)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LMU at:  
http://image.llnl.gov  
Plate: L14M9767 row: p column: 20  
High quality sequence start: 17  
High quality sequence stop: 724.  
Location/Qualifiers  
1. 918  
/organism="Mus musculus"  
/strain="FVB/N"  
/db\_xref="taxon:10090"  
/clone="IMAGE:4206595"  
/clone\_id="NCI\_CGAP\_C024"  
/lab\_host="DH10B (TI phase-resistant)"

/note="Organ: colon; Vector: pCMV-Sport6; Site:1: NotI;  
Site:2: SalI; Cloned unidirectionally. Primer: 011go dt.  
Average insert size 1.6 kb. Constructed by Life  
Technologies. Note: this is a NCI\_CGAP Library."

BASE COUNT 153 a 292 c 268 g 204 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:	1,31e-76	Length:	918
Score:	913.00	Matches:	191
Percent Similarity:	91.74%	Conservative:	9
Best Local Similarity:	87.61%	Mismatches:	18
Query Match:	63.23%	Indels:	3
DB:	12	Gaps:	0

US-09-245-198a-4 (1-284) x BF577781 (1-918)

QY 67 LeuAlaCysLeuGlyLeuLeuAlaValValSerLeuGlySerArgAlaSerLeuSer 86  
|||||  
Db 2 CTGGCTGCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 61  
QY 87 AlaGlnGluProAlaGlnGluGluLeuValAlaGluGluAspGlnAspProSerGluLeu 106  
|||||  
Db 62 GCCCAGGAGCCTTCTCAGAGAGAGCTGACAGAGAGACCGCGGAGACCCCTGAACTG 121  
QY 107 AsnProGlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeuValArgPro 126  
|||||  
Db 122 AATCCCGACAG 181  
QY 127 ArgArgSerAlaProLysGlyArgGlyThrArgAlaArgAlaIleAlaHisTyr 146  
|||||  
Db 182 CGAAGAAATGCTCTCAAGAGCGGAGAGCGGCGCT -CGCCAGCATTTGACGCCATAT 240  
QY 147 GluValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSer 166  
|||||  
Db 241 GAGGTTCATCTCCGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
QY 167 GlyTrrPgluIleAlaArgIleAsnSerSerSerProLeuArgTyrAsnArgGlnIleGly 186  
|||||  
Db 301 GCGTGGAGAGAGACCAATCAACAGCTCCGCTGCTGCTGCTGCTGCTGCTGCTG 360  
QY 187 GluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGlu 206  
|||||  
Db 361 GAATTACAGCATCAGAGGCGGCTCTACTACCTGCTGCTGCTGCTGCTGCTGCTG 420  
QY 207 GlyLysAlaValTyrLeuLysLeuAspLeuValAspGlyValLeuAlaLeuArgCys 226  
|||||  
Db 421 GGAAGGCTGTCTACTGACCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
QY 227 LeuGluGluPheSerAlaThrAlaIleSerSerLeuGlyProGlnLeuArgLeuGln 246  
|||||  
Db 481 CTGGAGAAATTTCTACGCCACAGACAGAGCTCTCTGGGCGCCACCTCGTTGGCCAG 540  
QY 247 ValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProTrr 266  
|||||  
Db 541 GTGTCTGGCGTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 599  
QY 267 AlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 284  
|||||  
Db 600 GCTCATCTTAAGGCTGC -CCCTTCTACTACTTGGACTCTTCAAGTTTCAC 652

RESULT 7  
BQ707185  
LOCUS 948 bp mRNA linear EST 16-JUL-2002  
DEFINITION AGENCOURT\_8353983 NIH\_MGC\_113 Homo sapiens cDNA clone IMAGE:6278608  
5', mRNA sequence.  
ACCESSION BQ707185  
VERSION BQ707185.1 GI:21846084  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 948)  
 AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Dr. Mark Watson  
 cDNA Library Preparation: Rubin Laboratory  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1CM2466 row: n column: 17  
 High quality sequence start: 24  
 High quality sequence stop: 550.  
 Location/Qualifiers

## FEATURES

source  
 1..948  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="NIH-MGC\_113"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: spleen; Vector: pOTB7; Site\_1: XhoI; Site\_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."  
 BASE COUNT 173 a 334 c 252 g 187 t 2 others  
 ORIGIN

## Alignment Scores:

Pred. No.: 2 95e-76 Length: 948  
 Score: 909.50 Matches: 179  
 Percent Similarity: 98.358 Conservative: 0  
 Best Local Similarity: 98.358 Mismatches: 2  
 Query Match: 62.984 Indels: 2  
 DB: 14 Gaps: 1

US-09-245-198a-4 (1-284) x BQ707185 (1-948)

QY 103 ProSeRgluLeuAsnProGlnThrGluSerGlnAspProAlaProPheLeuAsnArg 122  
 |||||||  
 Db 3 CCCTCGGAAGTGAATCCCGACAGAGAAAGACAGAGATCTCGCCTTCCTGAACCGA 62  
 QY 123 LeuValArgProArgArgSerAlaProLysGlyArgLysThrArgAlaArgAlaIle 142  
 |||||||  
 Db 63 CTGATTGCGGCTCGAAGAACTGCACCTAAAGCGCGGAAACACGCGCTCGAAGCGATC 122  
 QY 143 AlaAlaHisArgLysValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAsp 152  
 |||||||  
 Db 123 GCGAGCCATTATAGAGTTCAATCCACGACCTGCGACAGACGAGC---GCGAGGTGGAC 178  
 QY 163 GlyThrValSerGlyTyrGluGlnAlaArgLysAsnSerSerSerProLeuArgTyrAsn 182  
 |||||||  
 Db 179 GGGACAGTAGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 238  
 QY 183 ArgGlnIleGlyLysPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnVal 202  
 |||||||  
 Db 239 CGCCACATCGGGAGATTATAGTCACCCGGGCTGGGCTCTACTACTGTCTCAGGTG 258  
 QY 203 HisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGlyValLeu 222  
 |||||||  
 Db 299 CACTTGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 358  
 QY 223 AlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeu 242  
 |||||||  
 Db 359 GCCCTCGCGCTCGGAGGAGATTCACGACCTGCGGCGGAGTTCCCTCGGCGGCCAGCTC 418  
 QY 243 ArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArg 262

Db 419 CGCCTCGCAGAGTGTCTGGGCTGTGGCCCTCGGCGCAGGAGTCCCTCGGAGATCCGC 478  
 QY 263 ThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGln 282  
 |||||||  
 Db 479 ACCCTCCCTCGGCGCATCTCAAGGCTGCCCCCTTCTCAGCTTCTGAGCTTTCAG 538

QY 283 ValHis 284  
 |||||||  
 Db 539 GTTCAC 544

## RESULT 8

B0884231 940 bp mRNA linear EST 16-AUG-2002  
 LOCUS B0884231  
 DEFINITION AGENCOURT 8682031 lupskl.scfatic.nerve Homo sapiens cDNA clone  
 IMAGE:6197488 5', mRNA sequence.

ACCESSION B0884231 GI:22276239  
 VERSION B0884231.1  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 940)  
 AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: Dr. James R. Lupski  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: L1AM13607 row: 1 column: 17  
 High quality sequence stop: 453.  
 Location/Qualifiers

## FEATURES

source  
 1..940  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="IMAGE:6197488"  
 /clone\_lib="Lupskl.scfatic.nerve"  
 /sex="male"  
 /tissue\_type="sciatic nerve"  
 /dev\_stage="adult, 70 yr"  
 /lab\_host="DH10B"  
 /note="Vector: PCMV-SPORT6 (Life Technologies); Site\_1: NotI; Site\_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCAGCGCTCG-3' and 5'-GACTGTTCTAGTACGCGAGCGGCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."  
 BASE COUNT 170 a 325 c 255 g 190 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 9.73e-76 Length: 940  
 Score: 904.00 Matches: 183  
 Percent Similarity: 97.878 Conservative: 1  
 Best Local Similarity: 97.348 Mismatches: 0  
 Query Match: 62.608 Indels: 4  
 DB: 14 Gaps: 0

US-09-245-198a-4 (1-284) x B0884231 (1-940)

QY 101 GlnAspProSeRgluLeuAsnProGlnThrGluSerGlnAspProAlaProPheLeu 120  
 |||||||

Db 1 CAGGACCCGTGCGAATGATCCACAGACAGAAAGCCAGATCTCGCCTTCTCG 60  
QY 121 AsnArgLeuValArgProArgSerAlaProlysglyArgAlaArg 140  
Db 61 AACCGACTGTCGGCTCCGAGAGTCACCTAAAGCCGGAACACGGGCTCGAAGA 120  
QY 141 AAlaIleAlaIleHsTYrGluValHsProArgProGlyGlnAspGlyAlaGlnAlaGly 160  
Db 121 GCGATCCGACCCCATTTATGAAGTTCATCCAGCAGCTGACAGAGAGCCAGCAGGT 180  
QY 161 ValAspGlyThrValSerGlyTTPGluGlnAlaArgIleAsnSerSerProLeuArg 180  
Db 181 GTGACGCGACAGTAGAGTGGCTGGAGAGACCAAGATCAACAGCTCCAGCCCTCGCC 240  
QY 181 TyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200  
Db 241 TACAAACCCGACATCGGGAGATTATAGTACACCGGGCTGGGCTTACTACTGACTGT 300  
QY 201 GlnValHsPheAspGlyGlyLysAlaValTyrLeuLysLeuAspLeuValAspGly 220  
Db 301 CAGGTGCACTTTGATGAGAGGAGAGGCTGTCTACTGAAAGCTGAGCTGTGTGATGCT 360  
QY 221 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyPro 240  
Db 361 GTGCTGGCCCTGCTGCTGCTGAGAAATTCACGCACTGCGGCGCATTCCTCGGGGCC 420  
QY 241 GlnLeuArgLeuCys-GlnValSerGlyLeuLeuAlaLeuArgProGlySer-SerLeuA 260  
Db 421 CAGTCCGCTCTGCTCCCGAGGCTGTGGCTGTGGCCCTGGGCGAGGGCCCTCCCTCGC 480  
QY 260 rGileArgThrLeuProThrAlaHsLeuLysAlaAla-PropheLeuThrTyr-Phel 279  
Db 481 GGATCCGACCCCTCCCGGGGCTGAGCTGTGGCTGTGGCCCTGGGCGAGGGCCCTCCCTCGC 540  
QY 279 yLeuPheGlnValHs 284  
Db 541 ACTCTCCAGGCTCAC 556  
RESULT 9  
AW763237 561 bp mRNA linear EST 04-MAY-2000  
LOCUS ur70d09.y1 NCI-CGAP\_Mam3 Mus musculus cDNA clone IMAGE:3155633 5'  
DEFINITION similar to TR:054907 054907 TNF-RELATED WEAK INDUCER OF APOPTOSIS  
; mRNA sequence.  
ACCESSION AW763237  
VERSION AW763237.1 GI:7695174  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 561)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Other ESTs: ur70d09.x1  
Contact: Robert Strausberg, Ph.D.  
Email: cgaps-remail.nih.gov  
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMD)  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution Information can be  
found through the I.M.A.G.E. Consortium/LMD at:  
image.lnl.gov/image/html/lresources.shtml  
MG1:1058389  
Seq primer: -40RP from Gdbco  
High quality sequence stop: 433.  
Location/Qualifiers  
1..561  
/organism="Mus musculus"

/strain="129, C57BL/6J, FVB/N"  
/db.xref="taxon:10090"  
/clone="IMAGE:3155633"  
/clone\_lib="NCI-CGAP\_Mam3"  
/tissue\_type="tumor, gross tissue"  
/dev\_stage="10 months"  
/lab\_host="DH10B"  
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;  
Site:2: NotI. Cloned unidirectionally. Primer: Oligo dt.  
Library constructed by: Life Technologies. Investigators  
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH  
Reference for transgenic model: Xu et al., Nature Genetics  
22, 37-43 (1999)."  
BASE COUNT 108 a 158 c 194 g 100 t 1 others  
ORIGIN  
Alignment Scores:  
Pred. No.: 1.52e-68 Length: 561  
Score: 825.00 Matches: 161  
Percent Similarity: 90.91% Conservative: 9  
Best Local Similarity: 86.10% Mismatches: 17  
Query Match: 57.13% Indels: 0  
Gaps: 0  
US-09-245-198A-4 (1-284) x AW763237 (1-561)  
QY 41 SerGlnArgArgArgGlyArgArgGlyGluProGlyThrAlaLeuValProLeuAla 60  
Db 1 AGCCGAGGCGGAGGAGGCGCGGGGGAGAGCGGGGACCGGCTGCTGCTGGTG 60  
QY 61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaValSerLeuGly 80  
Db 61 CTGAGCCTGGGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 120  
QY 81 SerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuValAlaGluLys 100  
Db 121 AGCTGGGCAAGCTGTCTGCCAGAGAGCTTCTCAGAGAGAGCTGACAGAGAGAGCCG 180  
QY 101 GlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProheLeu 120  
Db 181 CGGAGGCCCTGAACTAAATCCCAAGACAGAGAAACCCAGATGTGATCTTCTTG 240  
QY 121 AsnArgLeuValArgProArgSerAlaProlysglyArgAlaArg 140  
Db 241 GACACACTAGTCCGCCCTCGAAGAGTCTCTTAAGCCCGAAGCCGCGCTCGCGA 300  
QY 141 AAlaIleAlaIleHsTYrGluValHsProArgProGlyGlnAspGlyAlaGlnAlaGly 160  
Db 301 GCTATTGCAAGCCCATTTATGAAGTTCATCCCTGGCGAGAGAGAGAGACACAGCAGGT 360  
QY 361 GTGATGGAGACAGTAGGCTGGGAGAGACCAAAATCAACAGCTCCAGCCCTTGCGC 420  
Db 181 TyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200  
Db 421 TACGACCCGACATTTGGGGAATNTACAGTCAACAGGCTGGCTTACTACTGACTGT 480  
QY 201 GlnValHsPheAspGlyGlyLysAlaValTyrLeuLysLeuAspLeuValAspGly 220  
Db 481 CAGGTGCACTTTGATGAGAGGAGAGGCTGTCTCTGAAAGCTGCTGTGTAACGCT 540  
QY 221 ValLeuAlaLeuArgCysLeu 227  
Db 541 GTGCTGGCCCTGCGCTGCTG 561  
RESULT 10  
BO671259 963 bp mRNA linear EST 15-JUL-2002  
LOCUS BO671259  
DEFINITION AGNCODRT\_8303564 NIH\_MGC\_102 Homo sapiens cDNA clone IMAGE:62747716  
5', mRNA sequence.  
ACCESSION BO671259  
VERSION BO671259.1 GI:21782093

KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE NIH-MGC http://mgi.nci.nih.gov/  
1 (bases 1 to 963)  
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: ATCC

CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM2456 row: 1 column: 13  
High quality sequence stop: 565.  
Location/Qualifiers

FEATURES  
source  
1. 963

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6274716"  
/clone\_1ib="NIH-MGC\_102"  
/issue\_type="epidermoid carcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: salivary gland; Vector: pORF7; Site\_1: XhoI; Site\_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG(6). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."

BASE COUNT 194 a 326 c 260 g 183 t  
ORIGIN

## Alignment Scores:

Pred. No.: 1,12e-66 Length: 963  
Score: 809.00 Matches: 160  
Percent Similarity: 99.388 Conservative: 0  
Best Local Similarity: 99.388 Mismatches: 1  
Query Match: 56.02% Indels: 1  
DB: 14 Gaps: 0

US-09-245-198a-4 (1-284) x B0671259 (1-963)

QY 124 ValArgProArgArgSerAlaProLysGlyArgLysThrArgAlaArgAlaIleAla 143  
DB 2 GTTCGGCGCTCGCAGAGTGCACCTAAAGCCGGGAAAAA-CGGGCTCGAAGAGCGATCGCA 60  
QY 144 AAlaHsTyrGluValHsProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGly 163  
DB 61 GCCCATTAAGACTTATCCACGACCTGGACAGGAGCGAGCGAGGTGGAGGG 120  
QY 164 ThrValSerGlyTyrPglGluAlaArgIleAsnSerSerSerProLeuArgTyrAsnArg 183  
DB 121 ACAGTAGAGTGGCTGGAGAGGAGCAGATCAACAGCTCCAGCCCTCGCCTCAACCCG 180  
QY 184 GlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHs 203  
DB 181 CAGATGGGGAGCTTTATAGCACCCGGGCTCTACTACTGTAAGTGTACAGTGCAC 240  
QY 204 PheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGlyValIleuAla 223  
DB 241 TTTGATGAGGGAGAGGCTGTCTACCTGAAGCTGGACTTGTGTGATGGTGTGGCC 300  
QY 224 LeuArgCysLeuGluGluPheSerAlaThrAlaIleSerSerLeuGlyProGlnLeuArg 243  
DB 301 CTCGGCTGGCTGAGGAATTCTCAGCCACTGGCGCCAGTTCCCTCGGCCCCCAGCTCCGC 360

QY 244 LeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThr 263  
DB 361 CTCCTGCAGAGTGTCTGGGCTGTGGCCCTCGGCGCAGGGTCTCCCTGGGATCCGAC 420  
QY 264 LeuProTAlaHsLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 283  
DB 421 CTCCTCGGGCCATCTCAAGGCTGCCCTTCCTCACTACTTGTGAGCTTCCAGGTT 480  
QY 284 Hs 284  
DB 481 CAC 483

RESULT 11  
LOCUS BM921213 1071 bp mRNA linear EST 12-MAR-2002  
DEFINITION AGENCOUNT 6633046 NIH-MGC\_115 Homo sapiens CDNA clone IMAGE:5752561  
5', mRNA sequence.  
ACCESSION BM921213  
VERSION BM921213.1 GI:19371592  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 1071)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLM12786 row: P column: 02  
High quality sequence stop: 656.  
Location/Qualifiers

## FEATURES

source

1. 1071  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5752561"  
/clone\_1ib="NIH-MGC\_115"  
/lab\_host="DH10B"  
/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH-MGC Library."

BASE COUNT 226 a 346 c 279 g 220 t  
ORIGIN

## Alignment Scores:

Pred. No.: 2.02e-62 Length: 1071  
Score: 765.00 Matches: 147  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 52.98% Indels: 0  
DB: 14 Gaps: 0

US-09-245-198a-4 (1-284) x BM921213 (1-1071)

QY 138 AlaArgAlaAlaIleAlaHsTyrGluValHsProArgProGlyGlnAspGlyAla 157  
DB 12 GCTCGAAGAGCGATCGACGCTTATGAGATTTCATCCAGACCTGGACAGGAGCG 71

QY	158	GINALAGIYALASGLYTHRValSerGIYTRPGIUGIUALAArgILEaNSerSerSer	177
Db	72	CAGCAGAGGTGTGGACGGGACAGTGAAGTGGCTGGGAGAGAACCCAGATATACAGCTCCAGC	131
QY	178	ProLeuAaGYrAsnAaRGlnIleGIyGluPheIleValThrAaGlaGlyLeuTYrTYr	197
Db	132	CCCTGCCTCAACACCCGACATCGGGAGATTATAGCACCCCGGCTGGCTCTACTAC	191
QY	198	LeuTYrCysGlnValHisPheAspGluGlyLysAlaValTYrLeuLysLeuAspLeu	217
Db	192	CTGACTCTCAGGTGCATTGTGATGAGGGGAAGGTGTCTACCTGAACCTGACTGTGTG	251
QY	218	ValAspGlyValLeuAlaLeuAaRGyCysLeuGluGluPheSerAlaThAlaAlaSerSer	237
Db	252	GTGATGTGTGTCGTGGCCCTGGCTGCTGGAGGAATCTCAGCAGACTCGGGCAGTTCC	311
QY	238	LeuGIYProGlnLeuAaRGLeuCYsGlnValSerGIYLeuLeuAlaLeuAaProGIYser	257
Db	312	CTCGGGCCCCAGCTCCGCTCTGCGCAGGTGTGTGGCTGTGGCCCTCGCCAGGGTCC	371
QY	258	SeLeuAaGlyIleAaGYThLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThTYr	277
Db	372	TCCCTGGCGGATCCGACACCTCCCTCGGGCCATCTCAAGGCTGCCCTTCTCACTAC	431
QY	278	PheGIYLeuPheGlnValHis	284
Db	432	TTCCGACTCTTCAGGTTTCAC	452
RESULT 12			
LOCUS	BO674188	951 bp	mRNA linear EST 15-JUL-2002
DEFINITION	AGENCOURT_8354100 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6275664		
ACCESSION	BO674188		
VERSION	BO674188.1	GI:21785022	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 951)		
TYPE	NIH-MGC http://mgs.nci.nih.gov/		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapds@mail.nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Rubin Laboratory		
	DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	Clone distribution by: Agencourt Bioscience Corporation		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		
	plate://LCM2459 row: d column: 01		
	High quality sequence stop: 335.		
FEATURES			
SOURCE	location/Qualifiers		
	1..951		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:6275664"		
	/clone_1id="NIH_MGC_102"		
	/tissue_type="epidermoid carcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
BASE COUNT	165 a	286 c	315 g 184 t 1 others
ORIGIN			

[illegible]

FEATURES	SOURCE
124 ValArgProGArgSseAlaProLySGlyArgLysThrArgAlaArgArgAlaIleAla	143
2 GTTGGCGCTCCAGAAAGTGCACACTTAAGCCGGAAACACGGGGCTCGAAGACGATCGCA	61
QY 144 AlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGly	163
Db 62 GCCCATTTGTGAAGTTTCATCCACGACCTGTGACGAGGAGCGCCAGCGCAGGTGTGGACGGG	121
QY 164 ThrValSerGlyTyrPrgLugLualAraArgIleAsnSerSerProLeuArgTyrAsnArg	183
Db 122 ACAGTAGATGGCTGGGAGGAGAACCGAAGATCAACACACTCCAGCCCTCTGGCGTACAAACCGC	181
QY 184 GlnIleGlyLupheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHis	203
Db 182 CAGATCGGGAGATTATATAGTCACCCGGGCTGGGCTCTACTACTGTACTGTACAGGTGCAC	241
QY 204 PheAspGluGlyLysAlaValTyrIleuLysLeuAspLeuValAspGlyValLeuAla	223
Db 242 TTTGATGTGGGGGAAAGCTGTCTACTCTGAAGCTGGCTGTGGTGTGTGTGGCC	301
QY 224 LeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArg	243
Db 302 CTGGCTCTCCGTGGAGGATTTCTCACCCACTGGGGCGCAGTCCCTGGGGGCCACGCTCCG	361
QY 244 LeuCysGlnValSerGlyLeuLeuAlaLeuArg-ProGlySerSerLeuArgIleArgTh	263
Db 362 CTCCTCCAGGTGTCTGGGCTGTGGCCCTCGCGGCGCAGGGTCTCTCCGTGGGATCCCCAC	421
QY 263 rLeuProTrpAlaHisLeuLysAlaAlaI-ProPheLeuThrTyrPheGlyLeuPhe	281
Db 422 CCTCCCGGGGCCCATTTAAAGGTGGCCCTTCTCTCACTATTTGGAAATTTT	477
RESULT 13	
LOCUS B1762908	
DEFINITION 785 bp mRNA linear EST 25-SEP-2001	
ACCESSION 60304/966F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:518168 5',	
VERSION B1762908	
KEYWORDS mRNA sequence.	
SOURCE B1762908.1 GI:15754486	
ORGANISM EST.	
human.	
Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
REFERENCE 1 (bases 1 to 785)	
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.	
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL Unpublished (1999)	
COMMENT Contact: Robert Strausberg, Ph.D.	
Email: cgabbs@remail.nih.gov	
Tissue Procurement: Life Technologies, Inc.	
cDNA Library Preparation: Life Technologies, Inc.	
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
DNA Sequencing by: Incyte Genomics, Inc.	
clone distribution: MGC clone distribution information can be	
found through the I.M.A.G.E. Consortium/LLNL at:	
http://image.llnl.gov	
Plate: L1AM1470 row: c column: 17	
High quality sequence stop: 647.	
Location/Qualifiers	
1..785	
/organism="Homo sapiens"	







gene 1. .1033  
/gene="tnsf12"  
<1. .433  
/gene="tnsf12"  
/note="data source:MGI, source key:MGI:1196259,  
evidence:ISS

putative  
tumor necrosis factor (ligand) superfamily, member 12"  
/codon\_start=2  
/protein\_id="BAB32249.1"  
/db\_xref="GI:12861641"  
/db\_xref="MGI:1196259"  
/translation="1AAHEVHPRGDGAQAGVDGVSQMEETKINSSPLRYDROI  
GEFTVRAGLVLYCOVHFDEGKAVYKLDLVNGVALRCLEEFSSATASSPGPOLR  
LCOVSGILPLRPSSSLRIRTLPAHLKAAPFLTYFGLFOVH"  
BASE COUNT 228 a 300 c 248 g 256 t 1 others  
ORIGIN

## Alignment Scores:

Pred. No.:	6.36e-57	Length:	1033
Score:	707.00	Matches:	135
Percent Similarity:	96.50%	Conservative:	3
Best Local Similarity:	94.41%	Mismatches:	5
Query Match:	48.96%	Indels:	0
DB:	11	Gaps:	0

US-09-245-198a-4 (1-284) x AK020909 (1-1033)

QY 142 ILAAlaAlaHisTyrGluValHisProArgProGlyValAspGlyAlaGluAlaGlyVal 161  
|||  
Db 2 ATTGCACGCCATTATGAGTTCATCCCTGGCCAGGACAGATGAGCACAAGCAGGTGTG 61  
|||  
QY 162 AspGlyThrValSerGlyTyrGluAlaArgIleAsnSerSerSerProLeuArgTyr 181  
|||  
Db 62 GATGGGACAGTGAAGTGGCTGGGAAGACCAAAATCAACAGCTCCAGCCTGGCTAC 121  
|||  
QY 182 AsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGln 201  
|||  
Db 122 GACCGCCAGATTGGGCAATTACAGTCATCAGGCTGGCTACTACTACTACTGTCAG 181  
|||  
QY 202 ValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGlyVal 221  
|||  
Db 182 GTGCACCTTATGAGGAAAGGCTGTCTACCTGAAGCTGAGCTGTGTAACGGGTGTG 241  
|||  
QY 222 LeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGln 241  
|||  
Db 242 CTGGCCCTGGCTGGCTGGAATTCACAGCACACAGCAAGCTCTCCTGGCCCCCAG 301  
|||  
QY 242 LeuArgLeuGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIle 261  
|||  
Db 302 CTCGGTTTGTGGCAGGTGTGGCGCTGGCCGCGGAGGTCCTCCCTTCGGATC 361  
|||  
QY 262 ArgThrLeuProTyrPalaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPhe 281  
|||  
Db 362 CGCACCTCCCTGGGCTCATCTTAAGGTGCCCTTCTTAACCTACTTTGGACTCTTT 421  
|||  
QY 282 GlnValHis 284  
|||  
Db 422 CAAATTCAC 430

Search completed: May 8, 2003, 04:22:52  
Job time : 1338.02 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:11:52 ; Search time 30.1297 Seconds  
(without alignments)  
1942.185 Million cell updates/sec

Title: US-09-245-198a-4  
Perfect score: 1444  
Sequence: 1 MSLLDFEISARRLPLPRSLG.....PMAHLKAPFLTYGLFQVH 284

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_protent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_virus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	8.0	325	5	Q9V5G2 drosophila
2	103	7.1	565	16	Q9KX66 streptomyce
3	101	7.0	643	16	Q9KX17 streptomyce
4	100.5	7.0	375	16	Q9RRI5 Q9RRI5 deinoococcus
5	100	6.9	850	11	Q9J1J5 mus musculu
6	99.5	6.9	378	16	Q9AAB9 Q9AAB9 caulobacter
7	99	6.9	614	4	Q8WU14 homo sapien
8	99	6.9	855	4	Q9NKK9
9	99	6.9	878	4	Q9UFU7
10	98	6.8	206	16	Q9S2W5
11	97.5	6.8	1560	4	Q9BJP2
12	97	6.7	408	10	Q8S5I5
13	97	6.7	926	4	Q9NYA0
14	97	6.7	1058	4	Q9Y4G2
15	96	6.6	629	10	Q8S1A6
16	95.5	6.6	536	4	Q9HB96

17	95.5	6.6	655	16	Q9FBR7	Q9FBR7 streptomyce
18	94.5	6.5	748	5	Q8RTY0	Q8RTY0 trypanosoma
19	94.5	6.5	1696	11	Q9WTR8	Q9WTR8 rattus norv
20	94	6.5	937	16	Q93UD1	Q93UD1 streptomyce
21	93.5	6.5	776	16	Q9RJ01	Q9RJ01 streptomyce
22	93.5	6.5	854	16	Q9F2P0	Q9F2P0 streptomyce
23	93.5	6.5	1038	10	Q9AS09	Q9AS09 oryza sativ
24	93	6.4	507	2	Q9RMT5	Q9RMT5 streptomyce
25	93	6.4	1043	10	Q93VC8	Q93VC8 oryza sativ
26	92.5	6.4	484	11	Q9CTA7	Q9CTA7 mus musculu
27	92.5	6.4	614	4	Q9NM41	Q9NM41 homo sapien
28	92.5	6.4	707	11	Q925Y8	Q925Y8 mus musculu
29	92.5	6.4	1511	13	Q91BE2	Q91BE2 fygu rubrip
30	92	6.4	351	12	Q85302	Q85302 orf virus.
31	92	6.4	566	16	Q70005	Q70005 streptomyce
32	92	6.4	845	4	Q96H68	Q96H68 homo sapien
33	92	6.4	977	4	Q8ET11	Q8ET11 homo sapien
34	91.5	6.3	212	16	Q86312	Q86312 mycobacteri
35	91.5	6.3	221	10	Q8RYZ6	Q8RYZ6 oryza sativ
36	91.5	6.3	664	16	Q9HYV9	Q9HYV9 pseudomonas
37	91.5	6.3	3169	5	Q917V8	Q917V8 drosophila
38	91	6.3	681	16	Q9H0Z6	Q9H0Z6 pseudomonas
39	91	6.3	1217	4	Q60336	Q60336 homo sapien
40	90.5	6.3	449	2	Q93HD5	Q93HD5 streptomyce
41	90.5	6.3	760	11	Q8VD14	Q8VD14 mus musculu
42	90.5	6.3	1100	2	Q937L3	Q937L3 deinoococcus
43	90.5	6.3	1329	16	Q9CD30	Q9CD30 mycobacteri
44	90	6.2	285	2	Q33629	Q33629 saccharopol
45	90	6.2	422	16	Q9LDB9	Q9LDB9 streptomyce

## ALIGNMENTS

RESULT 1

ID	Q9V5G2	PRELIMINARY:	PRT:	325 AA.
AC	Q9V5G2:	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	CG12919 protein.			
GN	CG12919.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BERKELEY;			
RX	MEDLINE-20196006; PubMed-10731132;			
RA	Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abdill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,			
RA	Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Bortova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,			
RA	Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabrielian A.E., Gary N.S., Gilbert W.M., Glasser K.,			
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hoskins D., Houston K.A., Howland T.J., Mei M.H., Ibegwan C.,			
RA	Jatani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			

RA Laakso P., Lehty A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Mekulov G., Milshina N.V., Moberly C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon R., Nuskern D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Port V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveloff J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003831; AAF5848.1;  
 DR FlyBase; FBgn0034483; CG12919.  
 DR InterPro; IPR000478; TNF\_Family.  
 DR SMART; SM00207; TNF\_1.  
 DR PROSITE; PS00251; TNF\_1; 1.  
 DR PROSITE; PS50049; TNF\_2; 1.  
 SQ SEQUENCE 325 AA; 36862 MW; 655CCB69694F1A3A CRC64;

Query Match 8.0%; Score 116; DB 5; Length 325;  
 Best Local Similarity 25.4%; Pred. No. 0.07;  
 Matches 57; Conservative 32; Mismatches 97; Indels 38; Gaps 10;

QY 82 RASASQAPQAEELVAE---EDDPSELNPQTESODPAPFLN-----LYPR 127  
 DB 118 RKSINADYRNEONTQNHTELQKSSNEATSKES--PALHHRHRRHRLHVRKG 175  
 QY 128 RSAPGRKTRARAIAYEYHPRPGDGAQAGVDTVSGVEARINSSPLRYNRQGE 187  
 DB 176 ESSLARSFEDSRP--AAHFHSSRRRHGSM-GYHGDMYIGNDRNRNYOG-HPOTRGV 231  
 QY 188 FIVRAGIYLYCOV-----HFDEKAYLYKLDLVGVLAIRCLLEESATPAASLGQOL 242  
 DB 232 LVTNTGTSLYVYADICYNNSHDQNFYF-----QSDTPRLQCLN---TVPTMPKRV 281  
 QY 243 RLCCVSGLLALPGSSLRIRTL--PMAHLKAAPLYEYGEFV 283  
 DB 282 HTCHTSGLIHLERNERHLKDIHNDRAVLAEGNNRSTFGIFKV 325

RESULT 2  
 Q9KY66 PRELIMINARY; PRT; 565 AA.  
 AC O9KY66;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Putative peptidase.  
 GN SC04798 OR SCD63A.09C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID:1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Brown S.P., Harris D.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 MEDLINE-97000351; PubMed-8843436;

RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,  
 RA Khashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2)."  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL356832; CAB92661.1;  
 DR InterPro: IPR002886; Peptidase\_M37.  
 DR Pfam: PF01551; Peptidase\_M37; 1.  
 SQ SEQUENCE 565 AA; 58070 MW; 7D0418D480C6A284 CRC64;

Query Match 7.1%; Score 103; DB 16; Length 565;  
 Best Local Similarity 21.2%; Pred. No. 1.7;  
 Matches 70; Conservative 33; Mismatches 99; Indels 128; Gaps 16;

QY 2 SLDFEISA---RLPLRSTGSRDCAVROAPAPMARSRORRGPEPTALL-- 56  
 DB 290 ALLDQDEATPDARIPARA-GSRAGARRRROP-----ALTTI 331  
 QY 57 -VPLALGL--GLALACGLLAVVSGRASASQAPQAE----- 94  
 DB 332 AVPSACVMSVAGIAAASVSLTG--DETEYAASPPDGNAPVPSAANKLDTQLT 389  
 QY 95 -LVAEED-----QDPSELNPQTESODPAPFLNRLVRRPSAPKGRTRARAIAYH 146  
 DB 390 SLAAGADDFADRASTRGTQRIDLKAQDAEKRAQAEARERLPRKFLPVKQGLSAYY 449  
 QY 147 EVHRRPGDG-----AAGVDTVSGVEARINSSPLRYNRQGE 187  
 DB 450 -----GQAGINMWSHTGIDPVLQGTYYAATDGYR-----TQNSAYGN 491  
 QY 188 FIVTRA--GLYLYCOVH-----FDEGKAYLYKLDLVGVLAIRCLLEESATPAAS 236  
 DB 492 MAITANDGTETWYCHLSYVPSGTYKAGDAI-----ATSGPSGN 533  
 QY 237 SLGQRLCQVSGLLALPGSSLRIRTLPW 266  
 DB 534 STGPHLH-----FEVVPAGGSSIDPLPW 556

RESULT 3  
 Q9KZ17 PRELIMINARY; PRT; 643 AA.  
 AC Q9KZ17;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Hypothetical protein SC02220.  
 GN SC02220 OR SC1087.15.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID:1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,



```

Db 601 LGSSTPALNGNPPRSPPEEPRRAGPTERRGNSGPKLIQALLRGTA-----LLA 654
QY 76 VVSIG-----SR---ASLSAQPAP-----EELVAEDDDPSLNPQTEESODPA 117
Db 655 SLGGRDLQPPGGLSRRGESPTAPPPQMPSPCPPELPSPLIRLSOTYDANHSPTPG 714
QY 118 PELNRL-----VRPRSPAPKGRKTRARRAIAHYVHPRGDGAQAGVDT 164
Db 715 PLLDLGVSPGSPAKSPRRETRGR-----TVSPPPGTSRSPAPGTPGT 758

RESULT 6
Q9AAB9 PRELIMINARY; PRT; 378 AA.
AC Q9AAB9;
AT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-MAR-2002 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Hyd family secretion protein.
GN CC0683.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_Taxid=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE-21173698; Pubmed-11259647;
RA Nielsen W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eriksen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocina I., Nelson W.C., Newton A., Stephens C., Phadke D.H.,
RA Deboy R.T., Dodson R.D., Durkin A.S., Gwinn M.L., Hatt D.H.,
RA Klonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005743; AAK2268.1;
DR TIGR; CC0683;
DR InterPro; IPR003997; RFXD.
DR PRINTS; PR01490; RXTXOXIND.
KW Complete proteome.
SQ SEQUENCE 378 AA; 39659 MW; 88B818C437C15C8C CRC64;

Query Match 6.9%; Score 99.5; DB 16; Length 378;
Best Local Similarity 25.1%; Pred. No. 2.1;
Matches 73; Conservative 29; Mismatches 106; Indels 83; Gaps 13;

QY 19 LGSRDGCAVRNQPAPAPARRSQRRGR---GEPGALLVPLALGLALACLGLILA 75
Db 36 LAAGGQGVANVHV-----REGDRVVGQPIATLTLSSALEGGDSFAVLSR--- 81
QY 76 VVSLGRSLASQEPAP---QEELVAEDDDPSLNPQTEESODPAFLNRPVRRS--AP 131
Db 82 --SLGAGSSAGSRAATQAALAESRQ-----LIGRRALNLELAARSLALQ 129
QY 132 KGRTRAR-----RAIAHYEVHPRGDGAQAGVDTGSGWEARINSSPLRYNQI 185
Db 130 TERLTLLARAEVRAETIAAGFLSPRE---LQARRSAELVAOEDASTLSSQALSYEROI 185
QY 186 GEFTVTRAGLYLVCOVHFDEGKRAVYLKLD---LYDGVTLARCLTEESATNAASLSRQL 242
Db 186 GEV-----DARLAAPITDQAARAELASTAGLEQ--QATQVEAGKRYV 227
QY 243 RLCOVSGILA-----LRPGSSLRTITPLPAHUKAPFL 275
Db 228 VVAIVAGRAVALPVEAGOTGVGAANVLTGDSALVAL- YASRAAGFV 277

RESULT 7
Q8WU14 PRELIMINARY; PRT; 614 AA.
ID Q8WU14

```

```

AC Q8WU14; Q9BR73;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Histone deacetylase 7A.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-COLON;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 339-614 FROM N.A.
RC TISSUE-B-CELL;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC020505; AAH20505.1;
DR EMBL; BC06453; AAH06453.1;
DR InterPro; IPR000286; His_deacetylase.
DR Pfam; PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
SQ SEQUENCE 614 AA; 66186 MW; 2B638D8C866B502 CRC64;

Query Match 6.9%; Score 99; DB 4; Length 614;
Best Local Similarity 24.2%; Pred. No. 4.2;
Matches 46; Conservative 21; Mismatches 67; Indels 56; Gaps 6;

QY 11 RLPLPSLSRGDCGAVR-----QAQPPAPMAR-----RSQRRGR- 47
Db 63 ROIISADLELDGAGGPGQVDDGLHRELGHGQPEARGPAPLQHPQVLLMEQRLRGL 122
QY 48 -RGPRTALVPLALGLIACGLLAVVSLGRSLASQEPAPQELVAEEDDDPSL 106
Db 123 PRGSGTDTVLPLAOGGHRPLS-----RAQSSPAAPASLSAPASARVLSSETPART 177
QY 107 NPQT-----EESQDP-----APFLNRLVPRRSAPKGRKTRAR 139
Db 178 LPFTTGILYDSVMLKHQCSGCDNSRHPDHAGRIQISVSLQERLRSQCCELRKASLE 237
QY 140 RAIAHYEVH 149
Db 238 ELQSVHSERH 247

RESULT 8
Q9NYK9 PRELIMINARY; PRT; 855 AA.
ID Q9NYK9
AC Q9NYK9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Histone deacetylase 7.
GN HDAC7.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-CERVICAL CARCINOMA;
RA Li S., Fische W., Verdin E., Walsh M.J.;
RT "A novel class II HDAC is associated with the transcriptional
RT homeodomain repressor CCAAT displacement protein.";
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF239243; AAF63491.1;
DR InterPro; IPR000286; His_deacetylase.
DR Pfam; PF00850; Hist_deacetyl; 1.
DR PRINTS; PR01270; HDASUPER.
SQ SEQUENCE 855 AA; 92363 MW; 3157426AE1913DB6 CRC64;

```

Query Match 6.9%; Score 99; DB 4; Length 855;  
 Best Local Similarity 24.2%; Pred. No. 6.2;  
 Matches 46; Conservative 21; Mismatches 67; Indels 56; Gaps 6;

OY 11 RLPLPRSLGSRDGA---OAPPPAAR-----RSQRRGR- 47  
 DB 361 ROIPEAELETFGGGPGVVDGLEHRELHGQPEARGPAPQHPVLLMEQRLAGRL 420  
 OY 48 -RGEFTALVPLALGTLALACGLLAVVSLGSRASISAOEPAOEELVAEEDDPSL 106  
 DB 421 PRGSTGDTVLLPLAOGGHRPLS-----RAQSSPAAPASISAEPPASQARVLSSETPART 475  
 OY 107 NPOT-----EESODP-----APFLNRLVPRRSAPKGRKTRAR 139  
 DB 476 LPFTTGLIYDSVMLKHQSCGSDNSRHPHAGRIQSIWRLQGRGLRSGCECLGRKASLE 535  
 OY 140 RAIAAHVEVH 149  
 DB 536 ELQSVHSERH 545

RESULT 9  
 O9UFU7 PRELIMINARY; PRT; 878 AA.  
 AC O9UFU7;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Hypothetical 94.3 kDa protein (Fragment).  
 GN DKF2P586J0917.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=UTERUS;  
 RA Koehler K., Beyer A., Mewes H.W., Gaassenhuber J., Wiemann S.;  
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AL117455; CAB55935.1;  
 DR InterPro: IPR000286; His\_deacetylase.  
 DR Pfam: PF00850; Hist\_deacetyl; 1.  
 DR PRINTS: PRO1270; HDASUPER.  
 DR Hypothetical Protein.  
 KW NON TER  
 FT  
 SO SEQUENCE 878 AA; 94330 MW; 26882E693960E10D CRC64;

Query Match 6.9%; Score 99; DB 4; Length 878;  
 Best Local Similarity 24.2%; Pred. No. 6.4;  
 Matches 46; Conservative 21; Mismatches 67; Indels 56; Gaps 6;

OY 11 RLPLPRSLGSRDGA---OAPPPAAR-----RSQRRGR- 47  
 DB 327 ROIPEAELETFGGGPGVVDGLEHRELHGQPEARGPAPQHPVLLMEQRLAGRL 386  
 OY 48 -RGEFTALVPLALGTLALACGLLAVVSLGSRASISAOEPAOEELVAEEDDPSL 106  
 DB 387 PRGSTGDTVLLPLAOGGHRPLS-----RAQSSPAAPASISAEPPASQARVLSSETPART 441  
 OY 107 NPOT-----EESODP-----APFLNRLVPRRSAPKGRKTRAR 139  
 DB 442 LPFTTGLIYDSVMLKHQSCGSDNSRHPHAGRIQSIWRLQGRGLRSGCECLGRKASLE 501  
 OY 140 RAIAAHVEVH 149  
 DB 502 ELQSVHSERH 511

RESULT 10  
 O9S2W5 PRELIMINARY; PRT; 206 AA.  
 AC O9S2W5;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Putative membrane protein.  
 GN SCO2091 OR SCA10.24C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 NX NCBI\_Taxid=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Saunders D.C., Harris D.;  
 RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kleser H.M., Denapalte D., Elchner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmid maps and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2)."  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL109663; CAB52000.1;  
 DR InterPro: IPR002965; P-rich\_extensn.  
 DR PRINTS: PRO1217; PRICHEXTENS.  
 SO SEQUENCE 206 AA; 20839 MW; 90131DF9AAB059D7 CRC64;

Query Match 6.8%; Score 98; DB 16; Length 206;  
 Best Local Similarity 27.3%; Pred. No. 1.3;  
 Matches 48; Conservative 18; Mismatches 64; Indels 46; Gaps 7;

OY 20 GSRDGAARQAOPAPMAARSQRRGR-----RGEFTA--LLVPLALGL 63  
 DB 6 GPRDGGKI-QTPPRGGRVSRKPELGRARLARLLPTGSGQAARAFVLLVVLGG 64  
 OY 64 GLALACGLLAVVSLGSRASISAOEPAOEELVAEEDDPSLNTQTESQAPAFNRL 123  
 DB 65 GL-----IGLVV-----NSALSSESSQQLDDLKQRTKELTDEQALORD 103  
 OY 124 VPRRSAPKGRKTRARRAIAAHVEVHPRPGDGAAGVGTGSGWEARINSSPL 179  
 DB 104 I-DAYSARRALQRRAR-----ELGWPVGGDPALPDGTYKGVSPAPAAATPL 151

RESULT 11  
 O96JP2 PRELIMINARY; PRT; 1560 AA.  
 AC O96JP2;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE KIA1783 protein (Fragment).  
 GN KIA1783.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=21245130; PubMed=11347906;  
 RT Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.:  
 "Prediction of the coding sequences of unidentified human genes. XX:  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro."  
 RL DNA Res. 8:85-95(2001).  
 DR EMBL: AB058686; BAB47412.1; -  
 DR InterPro: IPR000048; IQ\_region.  
 DR InterPro: IPR001609; myosin\_head.  
 DR InterPro: IPR000857; MYTH4.  
 DR InterPro: IPR001452; SH3.  
 DR Pfam: PF00612; IQ; 1.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF00784; MYTH4; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00096; IQ; 1.  
 DR PROSITE: PS50096; IQ; 1.  
 DR PROSITE: PS50002; SH3; 1.  
 PT NON-TER  
 SO SEQUENCE 1560 AA; 170146 MW; 6A61734C61EAC35A CRC64;

Query Match 6.8%; Score 97.5; DB 4; Length 1560;  
 Best local Similarity 25.5%; Pred. No. 17; Mismatches 117; Indels 67; Gaps 17;  
 Matches 75; Conservative 35;

QY 11 RLPLP-----RSLSRDGAVRQAQPPA---PMAARRSQRRC--RGEPTALLVPL 59  
 ID 11 RLPLP-----RSLSRDGAVRQAQPPA---PMAARRSQRRC--RGEPTALLVPL 59  
 AC 11 RLPLP-----RSLSRDGAVRQAQPPA---PMAARRSQRRC--RGEPTALLVPL 59  
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DE B2 gene partial cDNA, clone B2E.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=21245130; PubMed=11347906;  
 RT Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.:  
 "Prediction of the coding sequences of unidentified human genes. XX:  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro."  
 RL DNA Res. 8:85-95(2001).  
 DR EMBL: AB058686; BAB47412.1; -  
 DR InterPro: IPR000048; IQ\_region.  
 DR InterPro: IPR001609; myosin\_head.  
 DR InterPro: IPR000857; MYTH4.  
 DR InterPro: IPR001452; SH3.  
 DR Pfam: PF00612; IQ; 1.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF00784; MYTH4; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00096; IQ; 1.  
 DR PROSITE: PS50096; IQ; 1.  
 DR PROSITE: PS50002; SH3; 1.  
 PT NON-TER  
 SO SEQUENCE 1560 AA; 170146 MW; 6A61734C61EAC35A CRC64;

QY 11 RLPLP-----RSLSRDGAVRQAQPPA---PMAARRSQRRC--RGEPTALLVPL 59  
 ID 11 RLPLP-----RSLSRDGAVRQAQPPA---PMAARRSQRRC--RGEPTALLVPL 59  
 AC 11 RLPLP-----RSLSRDGAVRQAQPPA---PMAARRSQRRC--RGEPTALLVPL 59  
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DE B2 gene partial cDNA, clone B2E.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=21245130; PubMed=11347906;  
 RT Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.:  
 "Prediction of the coding sequences of unidentified human genes. XX:  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro."  
 RL DNA Res. 8:85-95(2001).  
 DR EMBL: AB058686; BAB47412.1; -  
 DR InterPro: IPR000048; IQ\_region.  
 DR InterPro: IPR001609; myosin\_head.  
 DR InterPro: IPR000857; MYTH4.  
 DR InterPro: IPR001452; SH3.  
 DR Pfam: PF00612; IQ; 1.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF00784; MYTH4; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00096; IQ; 1.  
 DR PROSITE: PS50096; IQ; 1.  
 DR PROSITE: PS50002; SH3; 1.  
 PT NON-TER  
 SO SEQUENCE 1560 AA; 170146 MW; 6A61734C61EAC35A CRC64;

RA Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;  
 RT "Genomic sequence for Oryza sativa, Nipprobase strain, clone  
 RT OSJNBA0061H20, from chromosome 10, complete sequence."  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC113337; NAM08840.1; -  
 KW Hypothetical protein.  
 SO SEQUENCE 408 AA; 41534 MW; CE4A6AB502840F45 CRC64;

Query Match 6.7%; Score 97; DB 10; Length 408;  
 Best local Similarity 26.5%; Pred. No. 3.8; Mismatches 84; Indels 66; Gaps 11;  
 Matches 60; Conservative 16;

QY 7 EISARRLPRLPSLSRDGAVRQAQPPA---ARRSQRRC--RGEPTALL 55  
 ID 7 EISARRLPRLPSLSRDGAVRQAQPPA---ARRSQRRC--RGEPTALL 55  
 AC 7 EISARRLPRLPSLSRDGAVRQAQPPA---ARRSQRRC--RGEPTALL 55  
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DE B2 gene partial cDNA, clone B2E.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=21245130; PubMed=11347906;  
 RT Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.:  
 "Prediction of the coding sequences of unidentified human genes. XX:  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro."  
 RL DNA Res. 8:85-95(2001).  
 DR EMBL: AB058686; BAB47412.1; -  
 DR InterPro: IPR000048; IQ\_region.  
 DR InterPro: IPR001609; myosin\_head.  
 DR InterPro: IPR000857; MYTH4.  
 DR InterPro: IPR001452; SH3.  
 DR Pfam: PF00612; IQ; 1.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF00784; MYTH4; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00096; IQ; 1.  
 DR PROSITE: PS50096; IQ; 1.  
 DR PROSITE: PS50002; SH3; 1.  
 PT NON-TER  
 SO SEQUENCE 1560 AA; 170146 MW; 6A61734C61EAC35A CRC64;

## RESULT 13

QYNTAO PRELIMINARY; PRT; 926 AA.

QY 11 RLPLP-----RSLSRDGAVRQAQPPA---PMAARRSQRRC--RGEPTALLVPL 59  
 ID 11 RLPLP-----RSLSRDGAVRQAQPPA---PMAARRSQRRC--RGEPTALLVPL 59  
 AC 11 RLPLP-----RSLSRDGAVRQAQPPA---PMAARRSQRRC--RGEPTALLVPL 59  
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DE B2 gene partial cDNA, clone B2E.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=21245130; PubMed=11347906;  
 RT Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.:  
 "Prediction of the coding sequences of unidentified human genes. XX:  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro."  
 RL DNA Res. 8:85-95(2001).  
 DR EMBL: AB058686; BAB47412.1; -  
 DR InterPro: IPR000048; IQ\_region.  
 DR InterPro: IPR001609; myosin\_head.  
 DR InterPro: IPR000857; MYTH4.  
 DR InterPro: IPR001452; SH3.  
 DR Pfam: PF00612; IQ; 1.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF00784; MYTH4; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00096; IQ; 1.  
 DR PROSITE: PS50096; IQ; 1.  
 DR PROSITE: PS50002; SH3; 1.  
 PT NON-TER  
 SO SEQUENCE 1560 AA; 170146 MW; 6A61734C61EAC35A CRC64;

Query Match 6.7%; Score 97; DB 4; Length 926;  
 Best local Similarity 24.2%; Pred. No. 10; Mismatches 125; Indels 56; Gaps 13;  
 Matches 69; Conservative 35;

QY 4 LDEISARRLPRLPSLSRDGAVRQAQPPA---ARRSQRRC--RGEPTALLVPL 63  
 ID 4 LDEISARRLPRLPSLSRDGAVRQAQPPA---ARRSQRRC--RGEPTALLVPL 63  
 AC 4 LDEISARRLPRLPSLSRDGAVRQAQPPA---ARRSQRRC--RGEPTALLVPL 63  
 DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
 DT 01-JUN-2002 (Tremblrel. 21, last sequence update)  
 DE B2 gene partial cDNA, clone B2E.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=21245130; PubMed=11347906;  
 RT Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.:  
 "Prediction of the coding sequences of unidentified human genes. XX:  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro."  
 RL DNA Res. 8:85-95(2001).  
 DR EMBL: AB058686; BAB47412.1; -  
 DR InterPro: IPR000048; IQ\_region.  
 DR InterPro: IPR001609; myosin\_head.  
 DR InterPro: IPR000857; MYTH4.  
 DR InterPro: IPR001452; SH3.  
 DR Pfam: PF00612; IQ; 1.  
 DR Pfam: PF00063; myosin\_head; 1.  
 DR Pfam: PF00784; MYTH4; 1.  
 DR Pfam: PF00018; SH3; 1.  
 DR SMART: SM00015; IQ; 1.  
 DR SMART: SM00096; IQ; 1.  
 DR PROSITE: PS50096; IQ; 1.  
 DR PROSITE: PS50002; SH3; 1.  
 PT NON-TER  
 SO SEQUENCE 1560 AA; 170146 MW; 6A61734C61EAC35A CRC64;



```

OY 117 -----APFLNR--LVPRRSAPKGRKTRARAIAMHYEVRPGDGAQAGVDGITS 167
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 337 YRGPGSPGILRHPSOEPRKNCISGALDQA-----CVSPGRQAAPS--QG 384
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 168 WEERARI-----NSSSPLRYNRQIGEIFYTRAGLY---YLCOVHEDEGAVYLLKLDL 217
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 385 HKSPRYVHRQMGSLNPRGLMKIG--TVERRGAMGWKELFCELSPLEFR-LVLSNE-- 439
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 218 VDGVLALRCLEEFSSATASSIGP-----QLRLCOVSGILALRPGS 257
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 440 -----EHTCVNCISLRCESVGPAPHSDFELVFSGKTLALRASS 479
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

RESULT 14
OY 09Y4G2 PRELIMINARY; PRT; 1058 AA.
AC 09Y4G2;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE KIAA0356 protein (Fragment).
GN KIAA0356.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA MEDLINE=9734984; PubMed=9205841;
RA Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
DR EMBL; AB002354; BAA20813.2;
DR InterPro: IPR002219; DAG_PE-bind.
DR InterPro: IPR000532; Glucagon.
DR InterPro: IPR001849;
DR InterPro: IPR004012; Run.
DR Pfam: PF00169; PH; 1.
DR Pfam: PF02759; RUN; 1.
DR PRINTS: PR00275; GLUCAGON.
DR SMART: SM00109; C1; 1.
DR SMART: SM00233; PH; 2.
DR PROSITE: PSS0081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE: PSS0003; PH_DOMAIN; 1.
FT NON_TER 1
SQ SEQUENCE 1058 AA; 117671 MW; CD1AD1869C8C9D1 CRC64;

Query Match 6.7%; Score 97; DB 4; Length 1058;
Best Local Similarity 24.2%; Pred. No. 12;
Matches 69; Conservative 35; Mismatches 125; Indels 56; Gaps 13;

```

```

DB 572 -----EHTCVNCISLRCESVGPAPHSDFELVFSGKTLALRASS 611
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

RESULT 15
OY 08S1A6 PRELIMINARY; PRT; 629 AA.
AC 08S1A6;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE B114B07.16 protein.
GN B114B07.16.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
RT clone:B114B07.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003334; BAB90006.1;
DR EMBL; 629 AA; 68639 MW; 724C08A73DCB785D CRC64;
SQ SEQUENCE 629 AA; 68639 MW; 724C08A73DCB785D CRC64;

Query Match 6.6%; Score 96; DB 10; Length 629;
Best Local Similarity 20.8%; Pred. No. 7.7;
Matches 67; Conservative 33; Mismatches 116; Indels 106; Gaps 11;

```

```

OY 12 RLPLRSLSGSDGAVRQAOPAPMAARRSQRRRREPEPT-----AL 55
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 107 RRPYRPPELRRAHYLRQAOPPRRERVRGRERRRRAVDPVDDRGDSRLRRRRRLS 166
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 56 LVPALGCG-----LALACL----- 70
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 167 LADTVAGLGSDVARRRGHLGCRSGGGGGGGLGAGSHATALLIPRGAPCRG 226
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 71 GLLIAVYSLGSRAS---LSAEPAGEELVAEEDDPSELNPOTESODPAPFLNRLVR 125
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 227 RGAALAPSLGRRRGCGRLGAGVGAEREADDEGRDPRGYEDRRGRGEPAEVALYV- 285
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 126 PRRSAPKGRKTRARAIAMHYEVRPGDGAQAGVDGTVSGMEARINSSPLRYNRQI 185
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 286 ---ERGGGGRVYRRRRHHWYRERERERESAAAGD-----RRRT 324
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 186 GEFTYTRAGLYLYCQVHEDEGKAVYLLKLDLLVDGLALRCLEFSATAASSIGPOLRLC 245
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 325 LEVAVT--CVHGLVCPHIAATARSL-----AACVLPYVAYVNAALASIGFG-RQC 372
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
OY 246 QVSGILALRPGSSLIRITLPMA 267
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
DB 373 VAG-----RCKSPYTPTRDMS 388
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Search completed: May 8, 2003, 02:21:27
Job time : 34.1297 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:07:47 ; Search time 10.6012 Seconds

(without alignments)  
1111.129 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 1444

Sequence: 1 MSLLDFEISARRLPRLSLG.....PMAHLKAPLFTYGLRQVH 284

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1268	87.8	249	TN12_HUMAN	O43508 homo sapien
2	1020	70.6	225	TN12_MOUSE	O54907 mus musculu
3	109.5	7.6	272	TNFS_CHICK	O91848 gallus gall
4	108.5	7.5	260	TNFS_CANPA	O91626 canis famli
5	106.5	7.4	254	TNFE_HUMAN	P41273 homo sapien
6	97	6.7	441	CG22_ANTHA	P34801 antirrhinum
7	95.5	6.6	201	TNFB_MACEU	O9xt48 macropus eu
8	94	6.5	280	TNFB_MACMU	O9my16 macaca mula
9	93.5	6.5	690	RHO_MICLU	P52154 micrococcus
10	93	6.4	310	Y497_MYCTU	O11162 mycobacteri
11	92	6.4	240	TN14_HUMAN	O43557 homo sapien
12	92	6.4	280	TNFB_CERTO	O9bdt1 cercocebus
13	92	6.4	902	NEC4_HUMAN	O914934 homo sapien
14	90.5	6.3	760	MLH1_MOUSE	O91k91 mus musculu
15	90.5	6.3	814	CADF_HUMAN	P55291 homo sapien
16	90	6.2	707	J1P1_MOUSE	O9wv19 mus musculu
17	89	6.2	280	MDCB_KLEPN	P71422 klebsiella
18	89	6.2	316	TN11_MOUSE	O35235 m tumor nec
19	88.5	6.1	278	TNFB_RAT	P36940 rattus norv
20	88.5	6.1	281	TNFB_HUMAN	P48023 homo sapien
21	87	6.0	197	TNFB_RABIT	P10154 oryctolagus
22	87	6.0	204	TNFB_BOVIN	O06600 bos taurus
23	87	6.0	291	TN10_MOUSE	P50592 mus musculu
24	86.5	6.0	250	TNFC_MACEU	O9xt47 macropus eu
25	86	6.0	139	TOFB_BACSU	P34467 bacillus su
26	86	6.0	205	TNFB_MARMO	O91m09 marmota mon
27	86	6.0	241	TN13_MOUSE	O64777 mus musculu
28	86	6.0	933	VG1B_HSYAL	O04457 herpesvirus
29	85.5	5.9	928	VG1B_HSYBP	P12471 bovine hepr
30	85.5	5.9	932	VG1B_HSYBC	P12640 bovine hepr
31	85	5.9	372	LKXB_MOUSE	O88609 mus musculu
32	85	5.9	379	LKXB_HUMAN	O60663 homo sapien
33	84.5	5.9	401	AROC_MYCTU	P95013 mycobacteri

34	84	5.8	284	1	HX1L_HUMAN	O43763 homo sapien
35	84	5.8	310	1	TNFC_MARMO	O91m10 marmota mon
36	84	5.8	575	1	MIS_PIG	P79295 sus scrofa
37	84	5.8	825	1	ICP0_HSV2H	P28284 herpes simp
38	83.5	5.8	416	1	RAGE_BOVIN	O28173 bos taurus
39	83.5	5.8	505	1	TUB_MOUSE	P50586 mus musculu
40	83.5	5.8	545	1	RTN2_HUMAN	O75298 homo sapien
41	83.5	5.8	885	1	VG1B_HSY2S	P24994 herpes simp
42	83	5.7	228	1	B10D_PSEAE	O91614 pseudomonas
43	83	5.7	653	1	APF1_MOUSE	O03157 mus musculu
44	83	5.7	936	1	FHL1_YEAST	P39521 saccharomyc
45	83	5.7	1217	1	AF4_MOUSE	O88573 mus musculu

## ALIGNMENTS

RESULT 1	ID	TN12_HUMAN	STANDARD:	PRT:	249 AA.
AC	O43508	OBWU27;			
DT	15-JUN-2002	(Rel. 41, Created)			
DT	15-JUN-2002	(Rel. 41, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TNFAK) (APO3 ligand).				
GN	TNFSF12 OR APO3L OR DR3LG.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND N-TERMINUS OF SOLUBLE FORM.				
RC	TISSUE-Tonsil, and Fetal liver;				
FX	MEDLINE=98070415; PubMed=9405449;				
RA	Chicheportriche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,				
RA	Hession C., Garcia I., Browning J.L.;				
RT	"TNFAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis."				
RL	J. Biol. Chem. 272:32401-32410(1997).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-Fetal kidney;				
FX	MEDLINE=9828355; PubMed=9560343;				
RA	Marsters S.A., Sheridan J.P., Pitti R.M., Brush J., Goddard A.,				
RA	Ashtkenazi A.;				
RT	"Identification of a ligand for the death-domain-containing receptor APO3."				
RL	Curr. Biol. 8:525-528(1998).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-Tonsil;				
RA	Strauberg R.;				
RL	Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.				
RN	[4]				
RP	FUNCTION.				
FX	PubMed=10085077;				
RA	Lynch C.N., Wang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S.R.;				
RT	"TNFAK induces angiogenesis and proliferation of endothelial cells."				
RL	J. Biol. Chem. 274:8455-8459(1999).				
CC	-1- FUNCTION: Binds to FN14 and possibly also to TNFSF12/APO3. Weak inducer of apoptosis in some cell types. Mediates NF-kappaB activation. May promote angiogenesis and the proliferation of endothelial cells.				
CC	-1- SUBUNIT: HOMOTRIMER (POTENTIAL).				
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein and secreted.				
CC	-1- TISSUE SPECIFICITY: Highly expressed in adult heart, pancreas, skeletal muscle, brain, colon, small intestine, lung, ovary, prostate, spleen, lymph node, appendix and peripheral blood lymphocytes. Low expression in kidney, testis, liver, placenta, thymus and bone marrow. Also detected in fetal kidney, liver, lung and brain.				
CC	-1- PM: The soluble form derives from the membrane form				

```

cc      by proteolytic processing.
cc      -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
cc      -1- CAUTION: Ref.3 sequence differs from that shown due to a
cc      frameshift in position 125.
cc      -----
cc      This Swiss-PROT entry is copyright. It is produced through a collaboration
cc      between the Swiss Institute of Bioinformatics and the EMBL outstation -
cc      the European Bioinformatics Institute. There are no restrictions on lists
cc      use by non-profit institutions as long as its content is in no way
cc      modified and this statement is not removed. Usage by and for commercial
cc      entities requires a license agreement (See http://www.isb-sib.ch/announce/
cc      or send an email to license@sib-sib.ch).
cc      -----
cc      EMBL; AF030099; AAC51923.1; -
cc      EMBL; AF055872; AAC39724.1; -
cc      EMBL; BC019047; AAH19047.1; ALT_FRAME.
cc      Genew; HGNC:11927; TNFSF12.
cc      MIM; 602695; -
cc      InterPro; IPR000478; TNF_family.
cc      Pfam; PF00229; TNF_1.
cc      SMART; SM00207; TNF_1.
cc      PROSITE; PS00251; TNF_1; FALSE_NEG.
cc      PROSITE; PS50049; TNF_2; 1.
cc      CycloKline; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.
cc      CHAIN 1 249
cc      TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
cc      TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
cc      MEMBER 12, MEMBRANE FORM.
cc      CYTOPLASMIC (POTENTIAL).
cc      SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
cc      (POTENTIAL).
cc      EXTRACELLULAR (POTENTIAL).
cc      CHAIN 94 249
cc      TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
cc      MEMBER 12, SECRETED FORM.
cc      CYTOPLASMIC (POTENTIAL).
cc      SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
cc      (POTENTIAL).
cc      EXTRACELLULAR (POTENTIAL).
cc      SITE 43 249
cc      CARBOHYD 139 139
cc      SEQUENCE 249 AA; 27216 MW; E66084361C28EBA CRC64;
cc
cc      Query Match 87.8%; Score 1268; DB 1; Length 249;
cc      Best Local Similarity 100.0%; Pred. No. 2e-93;
cc      Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

CC	Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Mus.
OX	NCBI_TaxID=10090.
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Peritoneal macrophage;
RC	MEDLINE=98070415; PubMed=9405449;
RX	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA	Hession C., Garcia I., Browning J.L.;
RT	"WEAK, a new secreted ligand in the tumor necrosis factor family that
RT	weakly induces apoptosis".
RL	J. Biol. Chem. 272:32401-32410(1997).
RN	[2]
RP	SEQUENCE OF 83-225 FROM N.A.
RC	STRAIN=C57BL/6J; TISSUE=Retina;
RC	MEDLINE=21085660; PubMed=11217851;
RX	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA	Atawa K., Iwawa M., Mishi K., Kiyosawa H., Kondo S., Yamana I.,
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA	Kadota K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA	Kuehl P., Lewis S., Matsuo Y., Nkaido I., Piesole G., Quackenbush J.,
RA	Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA	Blake J., Boffelli D., Bojunga N., Carinci P., de Bonaldo M.F.,
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Galbaldi M.,
RA	Guinleitch S., Hill D., Hofmann M., Hune D.A., Kamiya M., Lee N.H.,
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohsaki S.,
RA	Hayashizaki Y.;
RT	"Functional annotation of a full-length mouse cDNA collection.";
RL	Nature 409:685-690(2001).
RT	-1- FUNCTION: Binds to FN14 and possibly also to TNFRSF2/PO3. Weak
CC	inducer of apoptosis in some cell types. Promotes angiogenesis and
CC	the proliferation of endothelial cells. Mediates NF-kappa
CC	activation (by similarity).
CC	-1- SUBUNIT: Homotrimer (Potential).
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (by
CC	similarity).
CC	-1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC	-1- PIM: The soluble form is produced from the membrane form by
CC	proteolytic processing (by similarity).
CC	-1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/or_send_an_email_to_license@isb-sib.ch">http://www.isb-sib.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; AF030100; AAC53517.1; -
DR	EMBL; AK020909; BAB3249.1; -
DR	MGI; MGI:1196259; Tnfslf12.
DR	InterPro; IPR000478; TNF_family.
DR	Pfam; PF00229; TNF_1.
DR	SMART; SM00207; TNF_1.
DR	PROSITE; PS00251; TNF_1; FALSE-NEG.
DR	PROSITE; PS50049; TNF_2; 1.
KW	Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.
FT	NON_TER
FT	CHAIN
FT	<1
FT	225
FT	CHAIN
FT	70
FT	225
FT	TRANSMEM
FT	<1
FT	21
FT	DOMAIN
FT	22
FT	225
FT	SITE
FT	69
FT	70
FT	EXTRACELLULAR (POTENTIAL).
FT	CLEAVAGE (BY SIMILARITY).
FT	TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
FT	MEMBER 12, MEMBRANE FORM.
FT	TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
FT	MEMBER 12, SECRETED FORM (BY SIMILARITY).
FT	SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT	(POTENTIAL).
FT	EXTRACELLULAR (POTENTIAL).
FT	CLEAVAGE (BY SIMILARITY).



ID	TNFS9_HUMAN	STANDARD:	PRT;	254 AA.
AC	P41273;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Tumor necrosis factor ligand superfamily member 9 (4-1BB ligand) (4-1BBL).			
GN	TNFSF9.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94374434; PubMed=8088337;			
RA	Allderson M.R., Smith C.A., Tough T.W., Davis-Smith T., Armitage R.J., Falk B., Roux E., Baker E., Sutherland G.R., Din W.S., Goodwin R.G.;			
RT	Molecular and biological characterization of human 4-1BB and its ligand.";			
RL	Eur. J. Immunol. 24:2219-2227(1994).			
CC	-1- FUNCTION: Cytokine that binds to TNFSF9. Induces the proliferation of activated peripheral blood T cells. May have a role in activation-induced cell death (AICD). May play a role in cognate interactions between T cells and B cells/macrophages.			
CC	-1- SUBUNIT: HOMOTRIMER (POTENTIAL).			
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein.			
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN, PLACENTA, LUNG, SKELETAL			

RESULT 6	CG22.ANTMA	STANDARD:	PRT:	441 AA.
ID	CG22.ANTMA			
AC	P34801;			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	G2/mitotic-specific cyclin 2.			
OS	Antirrhinum majus (garden snapdragon).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
OC	Asteridae; euasterids I; Lamiales; Veroniceae; Antirrhinum.			
OX	NCBI_TaxID=4151;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=94148008; PubMed=8313906;			
RX	Robert P.R., Coen E.S., Murphy G.J.P., Doonan J.H.;			
RT	"Patterns of cell division revealed by transcriptional regulation of			
RA	genes during the cell cycle in plants."			

EMBO J. 13:616-624(1994).  
 RL FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G2/M  
 CC (MITOSIS) TRANSITION. G2/M CYCLIN ACCUMULATE STEADILY DURING G2  
 CC AND ARE ABRODLY DESTROYED AT MITOSIS.  
 CC -1- SUBUNIT: INTERACTS WITH THE CDC2 AND CK2 PROTEIN KINASES TO FORM  
 CC A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT  
 CC IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.  
 CC -1- DEVELOPMENTAL STAGE: ACCUMULATES STEADILY DURING G2 AND IS  
 CC ABRODLY DESTROYED AT MITOSIS.  
 CC -1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN AB SUBFAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X76123; CA53729.1; .  
 CC PIR: S41710; S41710.  
 CC HSSP: P30274; 1VIN.  
 CC InterPro: IPR004366; Cyclin.  
 CC InterPro: IPR004367; Cyclin\_Cterm.  
 CC Pfam: PF00134; cyclin; 1.  
 CC Pfam: PF02984; cyclin.C; 1.  
 CC SMART: SM00385; CYCLIN; 2.  
 CC PROSITE: PS00292; CYCLINS; 1.  
 CC CYCLIN: Cell cycle; Cell division; Mitosis.  
 CC KW CYCLIN 441 AA; 49205 MW; E6E4C037C9880A7 CRC64;  
 SO SEQUENCE

Query Match 6.7%; Score 97; DB 1; Length 441;  
 Best Local Similarity 24.0%; Pred. No. 1.6; Indels 74; Gaps 15;

Matches 70; Conservative 44; Mismatches 104; Indels 74; Gaps 15;  
 QY 25 GAVRQAQPPAPAPARRSOR-----RGRGEPGTALVPLALGLALACIGLL 73  
 DB 19 GAIKQK-----MAVEKKNRBALDIGNVYTVGVEGKALPVSRPTIRGE-----CAQGI 69  
 QY 74 -----LAVYISGSRASL-----AOEPAQELVAEDDOPSEINPOTEEQ 114  
 DB 70 ANAEAAEAENNNKNSLAVNAKGADGALPIKRAVARPVOKTKYSRQELIETISPTKEKK 129  
 QY 115 DEAPLNRLVPRRS-----APKGRKTRARAIAYEVPRPGQ-----DGAQAGVDGTV 165  
 DB 130 --APLEKEITGEKSLKKAPTLTSTLTARSAASY-VATKKEQIVDIDADVNDLAV 186  
 QY 166 SGWE-----ARINSSPLRY---NRQIGFIVTRAGLYLYLCQVHED---EGKAVYL 212  
 DB 187 VEYVEDMYKFKYSAENDSRPHDYMDOPETINEKM--RAILIDMLVQVHYKFKELSPETLYL 244  
 QY 213 KIDLVDVGLALRC-----LEFSATPAASLGPQLRCQVSGLLALRPGS 257  
 DB 245 TIN-IVDRYLASKTTSRRELQLLGMSMLIAKSYEIVNAPEVNDVICISDGS 295

## RESULT 7

TNFB\_MACEU  
 ID TNFB\_MACEU STANDARD: PRT: 201 AA.  
 AC 09XT48;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Lymphotoxin-alpha precursor (LT-alpha) (TNF-beta) (Tumor necrosis  
 DE factor ligand superfamily member 1).  
 GN LTA OR TNFSF1 OR TNFB.  
 OS Macropus eugenii (Tamar wallaby).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.  
 CX MCEL\_TaxID=9315;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20284948; PubMed=10826697;

Harrison G.A., Deane E.M.;  
 RA "cDNA cloning of lymphotoxin alpha (LT-alpha) from a marsupial,  
 RT Macropus eugenii.";  
 RL DNA Seq. 10:399-403(2000).  
 CC -1- FUNCTION: Cytokine that in its homotrimeric form binds to  
 CC TNFRSF1A/TNFR1, TNFRSF1B/TNFR and TNFRSF4/HVEM. In its  
 CC heterotrimeric form with LTB binds to TNFRSF3/LTR. Lymphotoxin is  
 CC produced by lymphocytes and cytotoxic for a wide range of tumor  
 CC cells in vitro and in vivo.  
 CC -1- SUBUNIT: Homotrimer, and heterotrimer of either two LTB and one  
 CC LTB subunits or (less prevalent) two LTB and one LTB subunits (By  
 CC similarity).  
 CC -1- SUBCELLULAR LOCATION: Secreted (homotrimer) and type II membrane  
 CC protein (heterotrimers) (By similarity).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AF119336; AAD41773.1; .  
 CC HSSP: P01374; 1TNR.  
 CC InterPro: IPR003636; TNF\_abc.  
 CC InterPro: IPR000478; TNF\_family.  
 CC Pfam: PF00229; TNF; 1.  
 CC PRINTS: PR01234; TNCRSISFCT.  
 CC Prodom: PD002012; TNF\_abc; 1.  
 CC SMART: SM00251; TNF; 1.  
 CC PROSITE: PS00251; TNF\_1; 1.  
 CC PROSITE: PS00409; TNF\_2; 1.  
 CC CYTOKINE; Glycoprotein; Cytotoxic; Signal.  
 CC KW SIGNAL 1 27  
 CC CHAIN 28 201  
 CC CARBOHYD 93 93  
 CC FT N-LINKED (GLYCANC. . .) (POTENTIAL).  
 SO SEQUENCE 201 AA; 21536 MW; 8CAC371CB5091627 CRC64;

Query Match 6.6%; Score 95.5; DB 1; Length 201;  
 Best Local Similarity 23.2%; Pred. No. 0.86;  
 Matches 44; Conservative 27; Mismatches 84; Indels 35; Gaps 9;  
 QY 107 NPOTESOPAPFNLRLVPRRSAPKGRTRARAI--AAHEVHPRPGDQAGVDGT 164  
 DB 30 NPDNSHSSPAP-----POTAQHSQKSLRETLKPAHL-----VDPVSQDSIH-- 75  
 QY 165 VSGWEARINSSP-LRYNRQI-GEFIVTRAGLYLYLCQVHDEKA-----VYKL 214  
 DB 76 ---W---RANTDHAFLRHGFSLSNSNLVPTSGLVFVYQVVFSGASCSEITPTLLYLISH 129  
 QY 215 DLIVDG---VLALRCLEEFSATPAASLGPQLRCQVSGLLALRPGSLRIRLPNAHKA 271  
 DB 130 EVLLFSKQYQVHVPPLLSAQKSYVSGTGPWMSYQGAVALITQGRDRLSTYDGVSHLQ 189  
 QY 272 APFLYFGLE 281  
 DB 190 SPSSVFFGAF 199

## RESULT 8

TNFB\_MACMU  
 ID TNFB\_MACMU STANDARD: PRT: 280 AA.  
 AC 09MTL6; Q9BDM5;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 6 (FAS antigen ligand)  
 DE (CD95L protein).  
 GN TNFSF6 OR FASL OR CD95L.  
 OS Macaca mulatta (rhesus macaque),  
 OS Macaca fascicularis (Crisp eating macaque) (Cynomolgus monkey), and  
 OS Macaca nemestrina (pig-tailed macaque).

Query Match	Score	DB 1	Length
50	280 AA; 31367 MW; F0B264D61A132EB4	CNC64;	280
FT	SEQUENCE		
FT	CONFLICT		
FT	CARBOHYD		
FT	CARBOHYD		
FT	259		
FT	249		
FT	183		
FT	201		
FT	129		
FT	45		
FT	4		
FT	DOMAIN		
FT	DOMAIN		
FT	102		
FT	280		
FT	69		
FT	4		
FT	DOMAIN		
FT	129		
FT	64		
FT	129		
FT	201		
FT	232		
FT	183		
FT	249		
FT	259		
FT	60		
FT	60		
FT	280 AA; 31367 MW; F0B264D61A132EB4	CNC64;	280
FT	SEQUENCE		
FT	CONFLICT		
FT	CARBOHYD		
FT	CARBOHYD		
FT	259		
FT	249		
FT	183		
FT	201		
FT	129		
FT	45		
FT	4		
FT	DOMAIN		
FT	DOMAIN		
FT	102		
FT	280		
FT	69		
FT	4		
FT	DOMAIN		
FT	129		
FT	64		
FT	129		
FT	201		
FT	232		
FT	183		
FT	249		
FT	259		
FT	60		
FT	60		
FT	280 AA; 31367 MW; F0B264D61A132EB4	CNC64;	280
FT	SEQUENCE		
FT	CONFLICT		
FT	CARBOHYD		
FT	CARBOHYD		
FT	259		
FT	249		
FT	183		
FT	201		
FT	129		
FT	45		
FT	4		
FT	DOMAIN		
FT	DOMAIN		
FT	102		
FT	280		
FT	69		
FT	4		
FT	DOMAIN		
FT	129		
FT	64		
FT	129		
FT	201		
FT	232		
FT	183		
FT	249		
FT	259		
FT	60		
FT	60		
FT	280 AA; 31367 MW; F0B264D61A132EB4	CNC64;	280
FT	SEQUENCE		
FT	CONFLICT		
FT	CARBOHYD		
FT	CARBOHYD		
FT	259		
FT	249		
FT	183		
FT	201		
FT	129		
FT	45		
FT	4		
FT	DOMAIN		
FT	DOMAIN		
FT	102		
FT	280		
FT	69		
FT	4		
FT	DOMAIN		
FT	129		
FT	64		
FT	129		
FT	201		
FT	232		
FT	183		
FT	249		
FT	259		
FT	60		
FT	60		
FT	280 AA; 31367 MW; F0B264D61A132EB4	CNC64;	280
FT	SEQUENCE		
FT	CONFLICT		
FT	CARBOHYD		
FT	CARBOHYD		
FT	259		
FT	249		
FT	183		
FT	201		
FT	129		
FT	45		
FT	4		
FT	DOMAIN		
FT	DOMAIN		
FT	102		
FT	280		
FT	69		
FT	4		
FT	DOMAIN		
FT	129		
FT	64		
FT	129		
FT	201		
FT	232		
FT	183		
FT			

```

Best Local Similarity: 20.8%; Pred. No. 1.6; Mismatches 62; Conservative 45; Mismatches 117; Indels 74; Gaps 14;

QY      13 LPLPRSLGSRDGGAVRQAOPAPMAARRSRRRGRRGEPTALLVPLAL----- 61
        || | : - | - | - | - | - | - | - | - | - | - | - | - | - | - |
Db       30 LPCPSTVYRRRGGQRPPPPPPLPL-----pppsplPLPLPLKKGNNHST 78
        || | : - | - | - | - | - | - | - | - | - | - | - | - | - | - |

QY      62 GLGLIACLIGLLAVLSVGSRASSAQSAPADELVAEDQPSELNPTTESQDPAPFLN 121
        || | : - | - | - | - | - | - | - | - | - | - | - | - | - | - |
Db       79 GLCLIVFEFVLVALGVSG-IGMFOLFHLOEL-----AELESTSKHTASSLER 128
        || | : - | - | - | - | - | - | - | - | - | - | - | - | - | - |

QY      122 RLVRPRRSAPKGRTRARARAIAHYEVPRRGGDAAGVDGYVSQWEEA-RINSSPLRL 180
        :: | : - | - | - | - | - | - | - | - | - | - | - | - | - | - |
Db       129 QIGRP----SPPEKKEQRK--VAHLTKPNRSRMPL-----WEDTGYILISVK 174
        :: | : - | - | - | - | - | - | - | - | - | - | - | - | - | - |

QY      181 YNRIOGEEIVTRAGLYLYLCVHFDEGKA-----VYLKLD-----LIYDGVIALLR 225
        -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -|
Db       175 YKK--GGGLIVNETGLEYFYSKVF- RQSGCINLPLSHKVYRNRSKYPDQLVMEEKMSYSY 231
        -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -|

QY      226 CLEFSATASSLIGPOLRLCOVSGILLRPOSSLRIPIRLPMHAHLKAAPFLYYGGLFOV 283
        -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -|
Db       232 CTTGOMAHASHSYLAGVENLTGSADHLX-----VVVSELSLVNEEESQ--TFEGLYKL 280
        -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -| -|

RESULT 9
RHO_MICLU STANDARD: PRT: 690 AA.
AC P52154;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Transcription termination factor rho.
GN RHO.
OS Micrococcus luteus (Micrococcus lysodeikticus).
OC Bacteria; Actinobacteriae; Actinobacteria (class); Actinobacteridae;
CC Actinomycetales; Micrococccinae; Micrococaceae; Micrococcus.
OX NCBI_TaxID=1270;
[1]
RN RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-5 AND 289-297.
RC STRAIN-EM;
RX MEDLINE=96132802; Pubmed=8557681;
RA Nowatzke W.L., Richardson J.P.;
RT "Characterization of an unusual Rho factor from the high G + C gram-positive bacterium Micrococcus luteus.";
RL J. Biol. Chem. 271:742-747(1996).
[2]
RN RP SEQUENCE OF 205-690 FROM N.A.
RC STRAIN-EM;
RX MEDLINE=94327472; Pubmed=8051015;
RA Opperman T., Richardson J.P.;
RT "Phylogenetic analysis of sequences from diverse bacteria with homology to the Escherichia coli rho gene.";
RL J. Bacteriol. 176:5033-5043(1994).
[3]
RN RP REVISION TO 500.
RC STRAIN-EM;
RA Nowatzke W.L.;
RL Submitted (Nov-1996) to the EMBL/GenBank/DDBJ databases.
CC -1- FUNCTION: FACILITATES TRANSCRIPTION TERMINATION BY A MECHANISM THAT INVOLVES RHO BINDING TO THE NASCENT RNA. ACTIVATION OF RHO'S RNA-DEPENDENT ATPASE ACTIVITY, AND RELEASE OF THE MRNA FROM THE DNA TEMPLATE. RNA-DEPENDENT NTTPASE WHICH UTILIZES ALL FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS WELL AS DMTF AS SUBSTRATES, BUT HAS A SIGNIFICANT LOWER ACTIVITY WITH CTP.
CC -1- SUBUNIT: HOMOHETERAMER (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
-----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/
```



or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch).

CC EMBL: L27277; AAB16671.1; -

DR HSSP; P03002; 1A63.

DR InterPro; IPR000194; ATPase\_a/bcentre.

DR InterPro; IPR002059; Cold\_shock.

DR InterPro; IPR004665; Term\_rho.

DR Pfam; PF00006; ATP-synt\_ab; 1.

DR SMART; SM00357; CSP; 1.

DR TIGRfams; TIGR00767; rho; 1.

KW Transcription termination; Helicase; ATP-binding; RNA-binding.

FT INT\_MET 0

FT DOMAIN 30 35 RNA-BINDING (RNP2) (BY SIMILARITY).

FT DOMAIN 318 321 RNA-BINDING (RNP1) (BY SIMILARITY).

FT NP\_BIND 441 448 ATP (POTENTIAL) (BY SIMILARITY).

FT CONFLICT 291 291 G->P (IN REF.; AA SEQUENCE).

SO SEQUENCE 690 AA; 75030 MW; F77C4C75E1B8998 CnC64;

Query Match

Best Local Similarity 25.9%; Score 93.5; DB 1; Length 690;

Matches 43; Conservative 19; Mismatches 63; Indels 41; Gaps 7;

QY 2 SLDPFISARLPPLRSLGSDGCAVROAPPAPMAARRSQRGRGEPGTALLVPLAL 61

DB 55 SVADRDAAERAAQAPAAFAETAPAAASSEDAAAP-AAERPARRRSRRAADATTS--APAAA 111

QY 62 GLGLALACIGLLLVSVLSRASLSAOPAEELVAEDODPELNPQTESODAPAPLNL 121

DB 112 QDQGPQA-----EAREAQTEQAPRE-----TASDQDRSGSGSEARDEGD----- 150

QY 122 RLVPRRSAPRGRKTRARRAIAAHYVPRGQDGAQGVDTGVS 167

DB 151 ---RPOSE-----RRSGRR-----RAGDDAQOGQDPRRSDG 179

RESULT 10

T497\_MTCUTU STANDARD; PRT; 310 AA.

AC Q11162;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Hypothetical protein RV0497.

GN RV0497 OR MT0517 OR MTCY2069. 23.

OS Mycobacterium tuberculosis.

OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae; Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI\_TaxID=1773;

ON [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H37RV;

RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F., Badcock R., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S., Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E., Taylor K., Whitehead S., Barrett L.G.;

RA "Elucidating the biology of Mycobacterium tuberculosis from the complete genome sequence."

RT Nature 393:537-544(1998).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CDC 1551 / Oshkosh;

RA Fleisshmann R.D., Alland D., Eisen J.A., Carpenter L., White O., Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L., Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishal W.;

RA "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains."

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: TO M.LEPRAE M2433.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).

CC EMBL: Z77162; CAB00923.1; -

DR EMBL: AE006952; AAK44740.1; -

DR TIGR; MT0517; -

DR Tuberculin; RV0497; -

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 231 251 POTENTIAL.

FT TRANSMEM 257 277 POTENTIAL.

FT TRANSMEM 286 306 POTENTIAL.

FT DOMAIN 33 39 POLY-ARG.

FT DOMAIN 197 202 POLY-ALA.

SO SEQUENCE 310 AA; 33092 MW; 4954027F694DF5C2 CnC64;

Query Match

Best Local Similarity 6.4%; Score 93; DB 1; Length 310;

Matches 80; Conservative 25; Mismatches 103; Indels 120; Gaps 18;

QY 14 PLPRSLGSRDGA---VRQAPPAPMAARRSQRGRG-----R 47

DB 6 PETESSGNROISVAVELARQGTGAP--ARRRRRRGSDATVAVELTGEIPIRDNHH 63

QY 48 RGEPTALLVPLATGLALACIGLLLVSVLSRASLSAOPAEELVAE----- 99

DB 64 ACPDAHASQSPANR-----VOYGEAAPSPAPVADQ-VAEPTTYVMS 109

QY 100 DQDPSELNPQTESODAPAPFLNLRVPRR---SAPRGKTRARRAI---AAHY----- 146

DB 110 QEPWRPKSPQDRRESGELSEYPRPLRHTSDRAPAGPGALMSDPVEHPDLWV 169

QY 147 -----EVPFRGQDS-----AQAGVDGIVSGHEARINSSS---PLKYN 182

DB 170 DVLDTVEGEAEAEVREARQAPGRGERHAAAAGTVDGGAARVARRALDVPYTLMR 229

QY 183 RQIGEIFYTR-----AGLYYLCOYHFDE---GRAVYIKLPL-----LVDSGLVR 225

DB 230 ---GALVYLQSTILAVAFGGLF-----IAFDLMRRNSIVLVSVNYILGLVSVRAVR 281

QY 226 CLEFSAT-----AASLAPOLRIQC 246

DB 282 KTEIASLTLINAVGALITLGP-LALLQ 308

RESULT 11

TN14\_HUMAN STANDARD; PRT; 240 AA.

AC Q43557; 075476; Q961D2; O8WV8;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Tumor necrosis factor ligand superfamily member 14 (Herpesvirus entry mediator-ligand) (HVEM-L).

GN TNFSF14 OR LIGHT OR HVEM-L.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

ON NCBI\_TaxID=9606;

OX [1]

RP SEQUENCE FROM N.A. (ISOFORM 1).

RX MEDLINE=98122340; PubMed=9462508;

RA Mauri D.N., Ebner R., Montgomery R.I., Kochel K.D., Cheung T.C., Yu G.-L., Ruben S., Murphy M., Eisenberg R.J., Cohen G.H., Spear P.G., Ware C.F.;

"LIGHT, a new member of the TNF superfamily, and lymphotoxin alpha are ligands for herpesvirus entry mediator.";  
 RL Immunity 8:21-30(1998).  
 [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.  
 RC TISSUE-LIVER;  
 RA HARPOLD-98438532; PubMed-9765287;  
 RA HARPOLD J.A., McDONNELL P.C., BRIGHAM-BURKE M., LYN S.D., MINTON J.,  
 TAN K.B., DEDE K., SPAMPANATO J., SILVERMAN C., HENSLEY P.,  
 DIPINZIO R., EMERY J.G., DEEN K., ELICHMAN C., CHABOT-FLETCHER M.,  
 TRUNEH A., YOUNG P.R.;  
 RA Herpesvirus entry mediator ligand (HVEM-L), a novel ligand for  
 HVEM/TR2, stimulates proliferation of T cells and inhibits HT29 cell  
 growth.";  
 RT J. Biol. Chem. 273:27548-27556(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND PROCESSING.  
 RA MEDLINE-21528948; PubMed-11673523;  
 RA GRANGER S.W., BUTROVICH K.D., HOSHMAND P., EDWARDS W.R., WARE C.F.;  
 RT "Genomic characterization of LIGHT reveals linkage to an immune  
 response locus on chromosome 19p13.3 and distinct isoforms generated  
 by alternate splicing or proteolysis.";  
 RT J. Immunol. 167:5122-5128(2001).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Brain;  
 RA Strausberg R.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytokine that binds to TNFRSF3/LTBR. Binding to the  
 decoy receptor TNFRSF6B modulates its effects. Activates NFkB,  
 stimulates the proliferation of T cells, and inhibits growth of  
 the adenocarcinoma HT-29. Acts as a receptor for Herpes simplex  
 virus.  
 CC -1- SUBUNIT: Homotrimer.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted  
 (isoform 1); Cytoplasmic (isoform 2).  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2/LIGHT  
 delta-TH; are produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE SPLEEN BUT ALSO  
 FOUND IN THE BRAIN. WEAKLY EXPRESSED IN PERIPHERAL LYMPHOID  
 TISSUES AND IN HEART, PLACENTA, LIVER, LUNG, APPENDIX, AND KIDNEY,  
 AND NO EXPRESSION SEEN IN FETAL TISSUES, ENDOCRINE GLANDS, OR  
 NONHEMATOPOIETIC TUMOR LINES.  
 CC -1- INDUCTION: UPREGULATED AFTER T-CELL ACTIVATION.  
 CC -1- PTM: N-glycosylated.  
 CC -1- PTM: The soluble form of isoform 1 derives from the membrane  
 form by proteolytic processing.  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -1- CAUTION: Ref.4 sequence differs from that shown due to a  
 frameshift in position 178.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF036581; AAC39563.1; -;  
 DR EMBL: AF064090; AAC25169.1; -;  
 DR EMBL: AY028261; AAK26160.1; -;  
 DR EMBL: BC018058; AAH18058.1; ALT\_FRAME.  
 DR HSSP: P01375; 4TSV.  
 DR Genew: HGNC:11930; TNFSF14.  
 DR MIM: 604520; -;  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR000478; TNF\_family.  
 DR Pfam: PF00229; TNF; 1.  
 DR ProDom: PD002012; TNF\_abc; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
 DR PROSITE: PS50049; TNF\_2; 1.

KW Cytokine; Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor;  
 KW Alternative splicing.  
 FT CHAIN 1 240  
 FT CHAIN 1 240  
 FT CHAIN 783 240  
 FT DOMAIN 1 37  
 FT TRANSMEM 38 58  
 FT DOMAIN 59 240  
 FT SITE 82 83  
 FT DISULFID 154 187  
 FT CARBOHYD 102 102  
 FT VARSPLIC 38 73  
 FT CONFLICT 120 120  
 FT CONFLICT 214 214  
 FT CONFLICT 214 214  
 SQ SEQUENCE 240 AA; 26351 MW; 49D0BF6E1390839 CRC64;  
 Query Match 6.4%; Score 92; DB 1; Length 240;  
 Best Local Similarity 23.7%; Pred. No. 2;  
 Matches 44; Conservative 19; Mismatches 57; Indels 66; Gaps 6;  
 QY 29 QAQPPAPMARRSORRRGRGEGPETAIVPLAIGLACGLGILLAVSLGSRASLSAQ 88  
 DB 16 QTDIPFTRLRSHRQSCSVARVGLGLL-LWGAGLAVOGWFLQLHMRG----- 66  
 QY 89 EPAQEEIYAEEDDPSELNPQTEESQDPAPFLNLYPRRSAPGRKTRARRAIAHYEV 148  
 DB 67 -----ENW-----TRPPDPAGSMWELIDERS-----HEV 92  
 QY 149 HPRPGDQAQAGDGVTSWEARINSSPLRNROI-----GEFVTRAGLY 196  
 DB 93 NPAAHLTGANSLSLNG-----SGGPLMETQGLAFLNGLSYHDALVYTKAGY 141  
 QY 197 YLYCOV 202  
 DB 142 YLYSKV 147  
 RESULT 12  
 ID TNF6\_CERVO STANDARD; PRT; 280 AA.  
 AC GQBDNI;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Tumor necrosis factor ligand superfamily member 6 (FAS antigen ligand)  
 DE (CD95L protein).  
 GN TNFRSF6 OR FASL OR CD95L.  
 OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).  
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Cercopithecus.  
 OX NCBI\_TaxID=9531;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Lymphocytes;  
 RA MEDLINE-21383618; PubMed-11491535;  
 RA VALLINGER F., BOSTIK P., MAYNE A.E., KING C.L., GENAIN C.P.,  
 RA Weiss W.R., Ansari A.A.;  
 RT "Cloning, sequencing, and homology analysis of nonhuman primate  
 Fas/FasL ligand and co-stimulatory molecules.";  
 RT Immunogenetics 53:315-328(2001).  
 CC -1- FUNCTION: Cytokine that binds to TNFRSF6/FAS, a receptor that  
 transduces the apoptotic signal into cells. May be involved in  
 cytotoxic T cell mediated apoptosis and in T cell development.  
 CC TNFRSF6/FAS-mediated apoptosis may have a role in the induction of  
 peripheral tolerance. In the antigen-stimulated suicide of mature  
 T cells, or both. Binding to the decoy receptor TNFRSF6B/DcR3  
 CC modulates its effects (by similarity).  
 CC -1- SUBUNIT: Homotrimer (Probable).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By  
 similarity).



FT DOMAIN 672 674 NUCLEAR LOCALIZATION SIGNAL.  
 SC SEQUENCE 902 AA; 95472 MW; E59F15F7647A7C6 CRC64;  
 Query Match 6.4%; Score 92; DB 1; Length 902;  
 Best Local Similarity 25.0%; Pred. No. 9;  
 Matches 48; Conservative 15; Mismatches 73; Indels 56; Gaps 8;  
 Q1 14 PLPRSLGSRDGAVALRQAQAPAPMAARRS-----QRRRGRRGEPGALLVPLALGLALAC 69  
 D1 236 PSPGGRPEDESWLLLSAPGTPASPSPGCKRRYSSTGSPSSA-----SPALSR 286  
 Q2 70 LGILLAVSISGRASLSAQEPDQPELVAEEDQDPSELNPQTESQDAPPLNRLVRRRS 129  
 D2 287 RG-----SLGEEGS-----EPPPPPL-PLARDGSGSPEDYGVGAPPAES 325  
 Q3 130 AP--KGRTRARRA-----AHYEVHPRGQCAAGVD---GVVSG 167  
 D3 326 IPQKTRTSSEQAVALLRSEBPASCKGLPLGAEESVAPGGRKVEYAGMDYLAVPSPLA 385  
 Q4 168 WEARINSSSPL 179  
 D4 386 WSKRIGHSPI 397  
 RESULT 14  
 MLH1\_MOUSE STANDARD; PRT; 760 AA.  
 AC Q9JK91; 062454;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE DNA mismatch repair protein Mlh1 (Mult. protein homolog 1).  
 GN MLH1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 FN [1]  
 RP SEQUENCE FROM N.A.  
 FC TISSUE=Testis;  
 RA Kumaran M., Rao M.R.S.;  
 RT "Cloning of the cDNA of the Mult. homolog, Mlh1 from mouse testis.";  
 IL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 IN [2]  
 IP SEQUENCE OF 1-151 FROM N.A.  
 IX MEDLINE=96270514; PubMed=8674118;  
 RA Edehmann W., Cohen P.E., Kane M., Lau K., Morrow B., Bennett S.,  
 RA Umar A., Kunkel T., Cattoretti G., Chaganti R., Pollard J.W.,  
 RA Kolodner R.D., Kucherlapati R.;  
 RT "Meiotic pachytene arrest in Mlh1-deficient mice.";  
 RL Cell 85:1125-1134(1996).  
 CC -1- FUNCTION: PROBABLY INVOLVED IN THE REPAIR OF MISMATCHES IN DNA.  
 CC -1- SUBUNIT: HETERODIMER OF MLH1 AND PMS2 OR MLH1 AND MLH3 (BY  
 CC SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MTH1/HEB FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL establishment -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF250844; AAF64514.1; -  
 CC EMBL: U60872; AAC52672.1; -  
 CC EMBL: U59881; AAC52672.1; JOINED.  
 CC EMBL: U59882; AAC52672.1; JOINED.  
 CC EMBL: U59883; AAC52672.1; JOINED.  
 CC EMBL: U59884; AAC52672.1; JOINED.  
 CC HSSP: P23367; 1BK1  
 DR MGD; MG1:101938; Mlh1  
 DR InterPro: IPR003594; ATPbind\_ATPase.

DR InterPro: IPR002099; DNA\_mis-repair.  
 DR Pfam: PF01119; DNA\_mis-repair; 1.  
 DR Pfam: PF02518; HAPbase\_C1.  
 DR TIGRfam: TIGR00585; mult; 1.  
 DR PROSITE: PS00058; DNA\_MISMATCH\_REPAIR\_1; 1.  
 DR DNA repair.  
 SQ SEQUENCE 760 AA; 84679 MW; 173C809372A29186 CRC64;  
 Query Match 6.3%; Score 90.5; DB 1; Length 760;  
 Best Local Similarity 22.7%; Pred. No. 9.8; Indels 81; Gaps 14;  
 Matches 70; Conservative 31; Mismatches 126;  
 Q1 13 LPLPRSLGSRDGAVALRQAQAPAPMAARRS-----QRRRGRRGEPGALLVPLALGLG 64  
 D1 436 LPAPAEAAASENLERSLMETSDAQAAPATSPSGSRKRRH--EDSDVEAVENMSGKE 493  
 Q2 65 LALAC-----LGILLAVSISGRASLSAQEPDQPELVAEEDQDPSELNPQ--TEESODPA 117  
 D2 494 MTACACYPERRRIINTSLSLQEEISERCHETLRE--ILRNHSFVGCVNPMALAHQTKL 551  
 Q3 118 PFLNRLVRRRSPAPKGRTRARRAIAHYEV-----HPRPGDGAQAGVDG 163  
 D3 552 YLLN-----TTKLSSELFYQILLYDPANFGYALSEPAPLFDLALALDS 596  
 Q4 164 TVSGWEARINSSSPLRYNRQIGFIVTRAGLYLYCQVHFDEGKAYTLKLLDLDG--- 220  
 D4 597 PESGWTDDGPKREGIAY--IYEFLLKKAEMLADYSVEIDE--EGMLIGLPLIDSYVP 652  
 Q5 221 -----VLAR-----CLEFESATAS--SLGPQ--LRICQVSGLLALRPGSS 258  
 D5 653 PLRGLPFLIRLATEVNMDEKBCFESLSKECAMFYIRQYILLESTLSGQSDMPGST 712  
 Q6 259 LRIRTLPM 266  
 D6 713 SK-----PW 716  
 RESULT 15  
 CADF\_HUMAN STANDARD; PRT; 814 AA.  
 ID CADF\_HUMAN  
 AC P55281;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Muscle-cadherin precursor (M-cadherin) (Cadherin-15) (Cadherin-14).  
 GN CDH15 OR CDH14 OR CDH3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 FN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=skeletal muscle;  
 RA Shihbata T., Shimoyama Y., Gotoh M., Hirohashi S.;  
 RT "Identification of human cadherin-14, a novel neurally specific type  
 RT II cadherin, by protein interaction cloning.";  
 RL J. Biol. Chem. 272:5236-5240(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Muscle;  
 RA Strausberg R.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE  
 CC SORTING OF HETEROGENEOUS CELL TYPES. M-CADHERIN IS PART OF THE  
 CC MYOCENT PROGRAM AND MAY PROVIDE A TRIGGER FOR TERMINAL MUSCLE  
 CC DIFFERENTIATION.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:13:22 ; Search time 15.6228 seconds  
(without alignments)  
1747.586 Million cell updates/sec

Title: US-09-245-198a-4  
1444  
Perfect score: 1 MSLLDFEISARRLLPRLPSIG.....PWAIHKAPFLTYFGLFQVH 284  
Sequence:

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. NO. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	106.5	7.4	254	2 I38427	4-1BB ligand - hum
2	104	7.2	1323	2 S27224	N-methyl-D-aspartate
3	100.5	7.0	375	2 A75264	hypothetical prote
4	99.5	6.9	378	2 H87333	HLyD family secret
5	99	6.9	878	2 T17245	hypothetical prote
6	98	6.8	206	2 T34961	probable membrane
7	97	6.7	441	2 S41710	mitosis-specific c
8	93.5	6.5	776	2 T36946	probable cation-tr
9	93.5	6.5	1657	2 T15838	hypothetical prote
10	93	6.4	310	2 D70745	hypothetical prote
11	92	6.4	351	2 B34768	ORF5 protein - Orf
12	92	6.4	566	2 T35203	probable two-compo
13	91.5	6.3	212	2 A70611	hypothetical prote
14	91.5	6.3	664	2 D83231	hypothetical prote
15	91	6.3	681	2 H83044	2,4-dienoyl-CoA re
16	91	6.3	1217	2 T00270	hypothetical prote
17	90.5	6.3	814	2 G02878	cadherin-15 precu
18	90.5	6.3	1329	2 D87226	conserved hypotet
19	90	6.2	660	2 T03038	probable inhibitor
20	89.5	6.2	886	2 S07132	hypothetical prote
21	88.5	6.1	210	2 D87394	hypothetical prote
22	88.5	6.1	278	2 A49266	fas ligand - rat
23	88.5	6.1	281	2 T38707	probable two-compo
24	88.5	6.1	998	2 G83022	conserved hypotet
25	88	6.1	531	2 C83153	hypothetical prote
26	87.5	6.1	492	2 A87471	tumor necrosis fac
27	87	6.0	197	1 JH0309	lymphotoxin - bov1
28	87	6.0	204	1 S24641	hypothetical prote
29	87	6.0	445	2 T30604	hypothetical prote

30	87	6.0	2124	2 H83357	probable non-ribos
31	86.5	6.0	399	2 F83633	hypothetical prote
32	86.5	6.0	439	2 B70629	hypothetical prote
33	86.5	6.0	755	2 B75346	probable competent
34	86.5	6.0	762	2 E98121	hypothetical prote
35	86.5	6.0	764	1 S14113	1-phosphatidylinos
36	86.5	6.0	810	2 D95256	Arp-dependent Clp
37	86.5	6.0	887	2 AG0535	ClpB-like protein
38	86	6.0	139	2 B69953	hypothetical prote
39	86	6.0	565	2 G98331	probable oligopept
40	86	6.0	568	2 E83325	probable chemotaxi
41	86	6.0	777	2 A87309	hypothetical prote
42	86	6.0	839	2 F75518	hypothetical prote
43	86	6.0	933	1 B48349	glycoprotein B pre
44	85.5	5.9	563	2 C98227	hypothetical prote
45	85.5	5.9	563	2 AE3059	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

138427

4-1BB ligand - human  
C:Species: Homo sapiens (man)  
C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 21-Jul-2000  
C:Accession: I38427

R.Alderson, M.R.; Smith, C.A.; Tough, T.W.; Davis-Smith, T.; Armitage, R.J.; Falk, B.  
Eur. J. Immunol. 24, 2219-2227, 1994

A:Title: Molecular and biological characterization of human 4-1BB and its ligand.  
A:Reference number: I38426; MUID:94374434; PMID:8088337

A:Accession: I38427

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-254 <RES>

A:Cross-references: EMBL:003398; NID:9571322; PIDN:AA5134.1; PID:9571323

Query Match 7.4%; Score 106.5; DB 2; Length 254;  
Best Local Similarity 26.5%; Pred. No. 0.32;  
Matches 74; Conservative 31; Mismatches 93; Indels 81; Gaps 13;

QY	32	PPAPMAARRSQRRRGRGEPALVPLALGIALACGLIAVSL-GSRASL-SAOE	89
DB	16	PPAP-----PAPACRVLP-WALVAGLLLLLLAAACAVFLACPMVAAGASPSAAS	67
QY	90	PAQELVAREDDPSSELNQTESQDPAFLNLTVPRRSAPGRKTRARBAIAEYEH	149
DB	68	PLURE-----GPELSP-----DDPAGLIDL-----RGMPAQLVAQNVLL-	102
QY	150	PRFGDQAQGVGVSGWEE--ARINSSPLRYNRQIGEFIVTRAGLYIYICQVHDE	206
DB	103	-----IDGPLSWSDPGLAGVSLTGLSTYEDRKELVNAAGYYVFFQ-----	146
QY	207	GAAYVTKLDLVDG-----VIALRCLEPSSATASSLGPQLRLCOVSG-----	249
DB	147	-----LELRVVAEGSGSVSLALHQLPLRSAGAAALLVTDVPPASSEARNASAFQOG	201
QY	250	-LLALRPSSLRIRTLPMWAIHKAPFL-----TYFGLEOV	283
DB	202	RLHLISAGQRLGVHTEARAHAMQLTGATVILGRV	240

##### RESULT 2

S27224

N-methyl-D-aspartate receptor epsilon-4 chain - mouse

C:Species: Mus musculus (house mouse)  
C:Date: 25-Feb-1994 #sequence\_revision 01-Sep-1995 #text\_change 17-Mar-1999  
C:Accession: S27224

R.Ikeda, K.; Nagasawa, M.; Mori, H.; Araki, K.; Sakimura, K.; Watanabe, M.; Inoue, Y.

FEBS Lett. 313, 34-38, 1992

A:Title: Cloning and expression of the epsilon-4 subunit of the NMDA receptor channel

A:Reference number: S27224; MUID:93050214; PMID:1385220

A:Accession: S27224

A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1323 <IKE>  
 A:Cross-references: EMBL:DJ2822  
 C:Superfamily: N-methyl-D-aspartate receptor 2D; glutamate receptor homology  
 C:Keywords: transmembrane protein  
 F:451-879/Domain: glutamate receptor homology <GRH>

Query Match 7.2%; Score 104; DB 2; Length 1323;  
 Best Local Similarity 23.7%; Pred. No. 3.2;  
 Matches 56; Conservative 21; Mismatches 97; Indels 62; Gaps 8;

QY 10 ARRLPLRSLSGSDGAVRQAPPAW-----AARRQRGRGEPGALLV----- 57  
 DB 903 AKRPPPPQPLSPAPVAPARPPGPAPVPRERAAADMRKAKG-TGPPGGAALADGPHRY 961  
 QY 58 -----PLALGLGLALAC-LGLLAAVSLGSRASLSAQPAAEELVAEEDDPSE----- 105  
 DB 962 YGPIEPQGLGEGARAAPRG--AAGRPPLSPPTTQPPQKPPSYFAIVREGEPAEPAGAF 1019  
 QY 106 -----LNPT-----EESQDPAPFLNRLVPRRS-----A 130  
 DB 1020 PGPSPPPAPPAAPAAAVGPPLCLRLAFEDSPAPASGRVLTTPRASRCWGARAARALGPR 1079  
 QY 131 PGRKTRARRAIAAHYEVHPRPGDGAQAGVDGTVGWEERAINSSPLRYNRQIG 186  
 DB 1080 PHRRVRRTAPPCCAYLDLEPSPDSSESLGASISGLGEPWFAFPYPIAERLG 1135

## RESULT 3

A75264  
 Hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000

C:Accession: A75264

R.White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; S.; Smith, H.O.; Venter, J.C.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.; McScience 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:2003696; PMID:10567266

A:Accession: A75264

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-375 <WHI>

A:Cross-references: GB:AE002081; GB:AE000513; NID:96460337; PIDN:AAFI2062.1; PID:9646034

C:Genetics:

A:Experimental source: strain R1

A:Gene: DR2516

A:Map position: 1

Query Match 7.0%; Score 100.5; DB 2; Length 375;  
 Best Local Similarity 24.0%; Pred. No. 1.5;  
 Matches 69; Conservative 27; Mismatches 84; Indels 107; Gaps 13;

QY 25 GAVRQADPPAPMAA-----RRSQRRGRGEGT--ALLVPLALGLALACIGLLAAV 77  
 DB 5 GRHRCQPTPGRAATARRLRGRRGRRGSPARRARRADLALALA-----LVA 58  
 QY 78 SLGSAASISAOE-PAQELVAEEDDPSELNPQTESQDPAPFLNRLVPRRSAPKGRKT 136  
 DB 59 DVVTRPTLPAGELPTLLDLAR-----QDLSDIDSP----- 89  
 QY 137 RARRAIAHYEVHPR-PGQDGAQG--VDGTVGWEERAINSSPLRYNRQIG-----F 188  
 DB 90 TDRILAVQARRATFRLPDPGAGFGHPISGLAGLE--RLSAASTRAMWRRFGORGSYLG 147  
 QY 189 IYTRAGIYLLY-----CQVHDEGKAVYKLLDLVDGVALALNCLEEF 231  
 DB 148 VVADADAQEVVELVAGLEADQWGEDRMRPHFPGRLRLH----- 188  
 QY 232 ATAASSIGPOLRLCOVS-----GLALRPSSSLRIRT 263

DB 189 ----PSATGEQTHSLVAPGPGPRDPWLPMQALITALSGSASRLFT 232

## RESULT 4

H87333

H1D family secretion protein [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C:Accession: H87333

R:Neuman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, R.; Leib, M.T.; Debroy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; K

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete Genome Sequence of Caulobacter crescentus.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: H87333

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-378 <STO>

A:Cross-references: GB:AE005673; NID:913421902; PIDN:AAK22668.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0683

Query Match 6.9%; Score 99.5; DB 2; Length 378;  
 Best Local Similarity 25.1%; Pred. No. 1.8;  
 Matches 73; Conservative 29; Mismatches 106; Indels 83; Gaps 13;

QY 19 LGSRDGAVRQADPPAPMAARRSORRRGR---GEGTALVPLALGLALACIGLLA 75  
 DB 36 LARQGGIYANVH-----REGDRVYVGQPIATLTSSALEGSDFAVLSR--- 81  
 QY 76 VVSLGSRASLSAQPAA--QEEVAEEDDPSELNPQTESQDPAPFLNRLVPRRS--AP 131  
 DB 82 --SLGQSSAASRAAATQALAAESRQ-----LTQRAALNRELAARAKSLAQ 129  
 QY 132 KGRKTRAR-----RAIAHYEVHPRPGDGAQAGVDGTVGWEERAINSSPLRYNRQI 185  
 DB 130 TRKTLTARAEVARAETIAAGFLSPRE---LQARRSALAVQQAESTLSSQALSYERQI 185  
 QY 186 GEYIYTRAGIYLLYCOVHEDEGKAVYKLLD---LVDGVALRCLEEFSTATAASLSGPOL 242  
 DB 186 GEV-----DARLAIPIIDLQAAAREASTLAGLEQ--QATQVEAQGRV 227  
 QY 243 RLCQVSGLLA-----LRPSSSLRIRTLPMHAKAPFL 275  
 DB 228 VVATVAGVAAALPVEAGQTVGVAIVLTTPDSALVALVET-VAPSRAGVF 277

## RESULT 5

T17245

Hypothetical protein DKFZp586J0917.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T17245

R:Kosher, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, September 1999

A:Reference number: Z18722

A:Accession: T17245

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-878 <KOE>

A:Cross-references: EMBL:AL117455

A:Experimental source: adult uterus; clone DKFZp586J0917

C:Genetics:

A>Note: DKFZp586J0917.1

Query Match 6.9%; Score 99; DB 2; Length 878;  
 Best Local Similarity 24.2%; Pred. No. 5.1;  
 Matches 46; Conservative 21; Mismatches 67; Indels 56; Gaps 6;

QY 11 RRLPLRSLSGSDGAVR-----QAQPPAPMAAR-----RSQRRRGR- 47  
 DB 327 RQIISADLETGGGPGQVVDGGLERELGHGPEARGPAPLQOHQPVLLMEODRLAGRL 386



Best Local Similarity 24.08; Pred. No. 3.4;  
Matches 70; Conservative 44; Mismatches 104; Indels 74; Gaps 10

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Mar-2000  
C:Accession: T15838

submitted to the EMBL Data Library, October 1995  
 A:Description: The sequence of *C. elegans* cosmid C54D2.  
 A:Reference number: Z18415

A:Accession: T15838

A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-1657 <MIN>

A:Cross-references: EMBL:037548; NID:g1017804; PID:g1017809; PIDN:AAV9201.1; CESP:C54D2

A:Gene: CESP:C54D2.5

A:Introns: 40/2; 67/3; 86/3; 121/2; 199/3; 230/2; 308/2; 334/2; 370/2; 439/3; 470/3; 513

C:Superfamily: sodium channel protein

Query Match

Best Local Similarity 23.2%; Score 93.5; DB 2; Length 1657;

Matches 63; Conservative 38; Mismatches 99; Indels 71; Gaps 12;

52 GTALIVPLAIGLALACLG-----LLAVVSLGSRAS-----LSAOEPAQELVAED 100

881 GMAQTNPMAALYFVALPTFGNYVLENLVAILVEGPESEKEERKOLEEDARKQAVEED 940

101 QDPSELNPOTEESSODPAPFLNRLVR-----RSAPKGRKTRARRAIAHYEVRPQ 154

941 ERKELELIIAKTSPA--FNNGVAPAECTCQRPSP--ESPRLLSANT--HSPER 994

155 DGAQAGVDGTVSGWEARINSSPL-----RYNRQIGEFIVTRAGLYLYCOV 202

995 KHS-ANLDAITD--KRLVLRNSAPFDRSPYSEGRDRLNHNASLIVPVANGVYPRQVR 1051

203 H-----FDEKAVYIKLDLYDGLALRCLE---EFSAATAASSLGPOLRICQV 247

1052 HSMKASQELKQALAEERKREAKQNTFVRKLKTKLHNTEFS----- 1095

248 SGLLALRPGSSLRITLPMALHKAAPLTYF 278

1096 --LPLMGPKNPLRIKCIQTQKWFDTYVLF 1124

RESULT 10

ID70745

hypothetical protein RV0497 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C:Accession: D70745

R:Coyle, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, T.; Fellwell, K.; Feltwell, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; M0ID:98295987; PMID:9634230

A:Accession: D70745

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-310 <COL>

A:Cross-references: GB:277162; GB:AL123456; NID:g3261606; PIDN:CAB00923.1; PID:e255036;

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV0497

Query Match

Best Local Similarity 24.4%; Score 93; DB 2; Length 310;

Matches 80; Conservative 25; Mismatches 103; Indels 120; Gaps 18;

14 PLPRSLGSRDGA---VROAOPAPMAARRSORRG-----R 47

6 PERESSNNROISVAELLARQGVGAP--ARRRRRRKDSALTYAELTGETPIIRDDHH 63

48 RGEFGALVPLAIGLALACLGILLAVVSLGSRASLSAOEPAQELVAE----- 99

64 AGPDNAHSQSPANR-----YQGEAAQSPAEPAED--VAEPEPTRYTYS 109

100 DQDSEINPOTESQDAPFLNRLVRP--SAPKGRKTRARRAI---AAHY----- 146

110 QPERBMPKSPQDRRESGPELSEYPRDLRHTSDRAPACPPSGAEHMSPDVEHYDPLW 169

147 -----EVHPRPGDG-----AQAGVCTVSGWEARINSS-----PLRYN 182

170 DVIDTEVEGEAEATEVEEADPGRGERRHAAAAAGTDEGGCAAEARVARRALDVPITLM 229

183 ROIGEFIVTR-----ACLYLYLCQVHDE-----GKAVYIKLDL-----LVGVLALR 225

230 ---GALVYLOSILAVAGALIF-----IAPDOLMRMSIYALVSVAVIIGLYVSPRAVR 281

226 CLEEFSAF-----AASSLGPOLRLCQ 246

282 KTEDIASTLIAVAVGALITLGP--LALLQ 308

RESULT 11

B34768

Ore6 protein - Orf virus (strain NZ2)

C:Species: Orf virus

C>Date: 23-Aug-1991 #sequence\_revision 23-Aug-1991 #text\_change 08-Oct-1999

C:Accession: B34768

R:Fraser, K.M.; Hill, D.F.; Mercer, A.A.; Robinson, A.J.

Virology 176, 379-389, 1990

A:Title: Sequence analysis of the inverted terminal repetition in the genome of the p

A:Reference number: A34768; M0ID:90266454; PMID:2129563

A:Accession: B34768

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-351 <FRA>

A:Cross-references: GB:M30023; EMBL:M37623; NID:g332561; PIDN:AAA46788.1; PID:g332564

Query Match

Best Local Similarity 23.0%; Score 92; DB 2; Length 351;

Matches 45; Conservative 18; Mismatches 73; Indels 60; Gaps 7;

14 PLPRSLGSRDGAAYROAPAPMAARRSQ--RRGRGEPSTALLVPLAIGLALACGL 72

211 PLPRRAR--GRORRGOPPPRRARRAQOPRRRARRAAG----- 247

73 LIAVSLGSRASLSAOEPAQELVAEDDOPSELNPOTESODPAPFLNRLVPRRSAPK 132

248 -----ARRGRGAPRQOQRPVORAAAQRRARQOROP 283

133 GKTRARRAIAHYE-VHPRGQDGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVT 191

284 PRVRARRARRARQQRARHOR--RRGRARRRCS-----RVYSKD-----SREVGWCD 332

192 RAGLYLYLCQVHDEG 207

333 KERRYIRVLLHFEHG 348

RESULT 12

T35203

probable two-component sensor - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 31-Jan-2000

C:Accession: T35203

R:Seeger, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, April 1998

A:Reference number: Z21571

A:Accession: T35203

A:Status: preliminary; translated from GB/EMBL/DBD

A:Molecule type: DNA

A:Residues: 1-566 <SEB>

A:Cross-references: EMBL:AL022374; PIDN:CAA18527.1; GSPDB:GN00070; SCOEDB:SC5B8.19C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC5B8.19C

C:Superfamily: two-component sensor histidine kinase; sensor histidine kinase homolog

Query Match

6.4%; Score 92; DB 2; Length 566;



```

0Y 105 ELNPQTEESODPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEHPRPGODGAQAGVDT 164
0b 549 --NSOPGGLREPA-----TRPARQV-----WLLQRRKGAPGAQL---GK 583
0Y 165 VSGMEARINSSSPLRN--RQIGFIVTRAGLYLYCQVHFDEGKAVYLLKD-----L 217
0b 584 TSGMVH-----RAHLRHNNAVRLG-----GVEYL---KIDE-RGLIRVDGVERMLE 626
0Y 218 VDGVL-----ALRCLEEFSAATASSLGPOLRL-----COVSGLL---ALRPGSSLRIR 262
0b 627 VDNVVICAGGEPLELQIROATES-----LRFHLIGGARVAGELDAKRAIRBGAMLAR 680

```

Search completed: May 8, 2003, 02:22:06  
Job time : 18.6228 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 8, 2003, 02:22:47 ; Search time 210.35 Seconds

(without alignments)  
3040.496 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 1444

Sequence: 1 MSLDFEISARRLPLPSISLG.....FWHLKAAPFLTYFGLFQVH 284

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO-POOL/US0924518/runat\_06052003\_170239\_9605/app.query.fasta\_1.846  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRAN=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptio -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US0924518 -ECGN\_1.1.323 -etunat\_06052003\_170239\_9605 -NCPU=6 -ICPU=3  
-NO\_XLPTX -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_101002: \*  
1: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1980.DAT: \*  
2: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1981.DAT: \*  
3: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1982.DAT: \*  
4: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1983.DAT: \*  
5: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1984.DAT: \*  
6: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1985.DAT: \*  
7: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1986.DAT: \*  
8: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1987.DAT: \*  
9: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1988.DAT: \*  
10: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1989.DAT: \*  
11: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1990.DAT: \*  
12: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1991.DAT: \*  
13: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1992.DAT: \*  
14: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1993.DAT: \*  
15: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1994.DAT: \*  
16: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1995.DAT: \*  
17: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1996.DAT: \*  
18: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1997.DAT: \*  
19: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1998.DAT: \*  
20: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA1999.DAT: \*  
21: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA2000.DAT: \*  
22: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA2001A.DAT: \*  
23: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA2001B.DAT: \*  
24: /SIDS2/gcgcdata/geneseq/geneseqn\_emb1/NA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1446	100.0	1373	19	AAV18600	Homo sapiens tumour
2	1406	97.4	1364	20	ABK34881	Human cDNA encoding
3	1377	95.4	1421	20	AAK56000	Human tumour necro
4	1366	94.6	1353	21	AAA49717	Human PRO207 cDNA
5	1366	94.6	1353	21	ABK40255	cDNA encoding huma
6	1265	87.6	1236	19	AAV47613	TNF related endoto
7	1265	87.6	1236	19	AAD04350	Human TRPA (TNF r
8	1066	73.8	898	22	AAK03964	Expression vector
9	1062	73.5	1030	20	AAK23424	Human TNFR3 DNA.
10	1020	70.6	1168	20	AAV18599	Mus musculus tumou
11	968	67.0	701	20	AAK23425	Mouse TNFR3 DNA.
12	127	8.8	2856	19	AAV24140	Homo sapiens BARD1
13	126.5	8.8	1660	21	AAA49205	Corn putative lecti
14	124.5	8.6	2832	20	AAV69289	Murine JNK-interac
15	124.5	8.6	154746	24	AAD25519	Human herpesvirus
16	124	8.6	3957	22	AAK09686	HSV-2 immediate ea
17	124	8.6	154746	24	AAD25519	Human herpesvirus
18	123.5	8.6	11705	19	AAV62160	HSV-2 strain SB5 C
19	123.5	8.6	117213	19	AAV62176	HSV-2 strain SB5 C
20	120.5	8.3	2828	21	AAA46674	cDNA of a human di
21	120.5	8.3	2828	24	ABK64503	Human benign prost
22	120.5	8.3	2828	24	ABK61804	Colon adenocarcino
23	117.5	8.1	3880	24	ABK67211	Thyroid cancer rel
24	117.5	8.1	3901	17	AAV42224	Human TATTA-binding
25	117.5	8.1	3901	18	AAV79598	TATTA-binding prote
26	117.5	8.1	3902	15	AAK07038	DNA encoding an ac
27	117	8.1	1030	20	AAV99230	Nucleotide sequenc
28	117	8.1	1030	21	AAV72514	Drosophila melanog
29	116	8.0	978	23	ABK21473	DNA encoding tumou
30	116	8.0	2148	24	ABK11680	Chicken beta-actin
31	115	8.0	1275	11	AAK01620	Expression vector of
32	115	8.0	1344	11	AAK01318	Hybrid promoter of
33	115	8.0	1345	11	AAK03062	Chicken beta-actin
34	115	8.0	1495	22	AAK84455	Human beta-actin
35	115	8.0	4211	22	AAH21792	Mouse Meg1/Grb10 n
36	115	8.0	5504	24	AAK43172	Rat expression pro
37	115	8.0	5565	24	AAK43171	Human expression p
38	115	8.0	5581	22	AAH41035	Hprt gene containi
39	115	8.0	5759	24	ABK49521	Positive selection
40	115	8.0	6148	24	ABK49520	pcGGS-hsEPOR2 rat
41	115	8.0	6256	24	AAK43173	Human PES1- $\alpha$ -pha
42	114.5	7.9	8033	22	AAK12739	Human PES1- $\alpha$ -pha
43	114.5	7.9	12700	19	AAV62133	HSV-2 strain SB5 C
44	113.5	7.9	3836	15	AAK053997	Vitamin D hydroxyl
45	113.5	7.9	44377	18	AAK78508	Platenolide syntha

## ALIGNMENTS

RESULT 1

ID AAV18600 standard; cDNA: 1373 BP.

AAV18600;

21-JUL-1998 (first entry)

Homo sapiens tumour necrosis factor related ligand (TNFR1) gene.

TNFR1; tumour necrosis factor related ligand; tnfr; treatment;

cancer; autoimmune disease; immune system; stimulation; suppression;

graft rejection; ds.

Homo sapiens. Location/Qualifiers

key CDS 1..852





CC to induce JNK/SAPK-dependent responses in mammalian cells.  
 XX  
 SQ Sequence 1421 BP; 281 A; 464 C; 404 G; 272 T; 0 other;

## Alignment Scores:

Pred. No.:	6 69e-91	Length:	1421
Score:	1377.00	Matches:	272
Percent Similarity:	97.84%	Conservative:	0
Best Local Similarity:	97.84%	Mismatches:	6
Query Match:	95.36%	Indels:	0
DB:	20	Gaps:	0

US-09-245-198a-4 (1-284) x AAA56000 (1-1421)

```

OY 7  G L I U L S E R A L A A R G A L E U P R O L E U P R O A R G S E R L E U G I Y S E R A R G A S P G I Y A L A 26
    |||||
DB 5  G A G A T C C C T G A C C T G C A C C G C G A T C C C G G G T C C C G G G A T G G G G G G C G 64

OY 27 V A L A R G L A A G I N P R O P R A L P R O M E T A L A A L A R G A S E R G I N A R G A R G L Y 46
    |||||
DB 65 G T G A G G C A G C A C A C C C C C C C A T G C C C C C G A G C C A G A G C G G A G G G G G 124

OY 47 A R G A R G L I G L U P R O G I Y T H R A L E U L E U V A L P R O L E U A L A L E U G I L E U G I L E U A 66
    |||||
DB 125 C G C G G G G G A G C C G G G C A C C G C T G T G C C C G C T G G C C T G G C C T G G C G 184

OY 67 L E U A L A C Y S L E U G I L E U L E U A L A V A L S E R L E U G I Y S E R A R G A L S E R L E U S E R 86
    |||||
DB 185 C T G C C T G C C T G C G C C T C T G C G C G T G T G C A G T T T G G G A G C C G G G C A T C G T G C 244

OY 87 A L A G I N G L U P R O A L A G I N G L U L E U V A L A G I L U G I U A S P G I N A S P P R O S E R G I L E U 106
    |||||
DB 245 G C C C A G A G C C T G C C C A G A G A G A C T G T G C A A G A G G A C C C T C G G A A C T G 304

OY 107 A S A P R O G I N T H R G L U S E R G I N A S P R O A L P R O P H E L A U A S A R G L E U V A L A R P R O 126
    |||||
DB 305 A A T C C C A C A C A A A A A A G C A G A T C T C C C C T T T C C T A A C C A C T A G T T C G G C C T 364

OY 127 A R G A R S E R A L A P R O L Y S G I Y A R G L Y S T H R A G L A A R G A L A L E A L A L A H I S T Y R 146
    |||||
DB 365 C G C A G A A G T G C A C T A A A G C G G A A A C A C G G G C T G A A G A G G A T G C A G C C A T T A T 424

OY 147 G L U V A L I S P R O A R G P R O G I Y G L I N A S P G I Y A L A G I N A G I Y A L A S P G I Y T H R V A L S E R 166
    |||||
DB 425 G A A G T T A T C C A C A C C T G C A G A C A G A C G A C G A G C G T G G A C G G A C A G T G A G T 484

OY 167 G L Y T R P S L U G I U A L A A R G I L A S N S E R S E R P R O L E U A R G Y R A S A R G L I L E G L Y 186
    |||||
DB 485 G G C T G G A G A G C A G C A G A T C A C A G C T C C A G C C C T C G C C T A C A C C G C A G A T C G G 544

OY 187 G L U P H E I L E V A L I T H R A R G A L A G L Y L E U T Y R T Y R L E U T Y R C Y S G I N V A L I S P H A S P G L U 206
    |||||
DB 545 G A G T T T A T A G T C A C C C G G G G G C T G T A C T A C T G T A C T G A G T G C A C T T T G A T G A G 604

OY 207 G L Y L Y S A L A V A L Y R L E U L Y S L E U A S P L E U V A L A S P G I Y V A L L E U A L A L E U A R G C Y S 226
    |||||
DB 605 G G G A A G G T G T A C C T G A C T G A C T G T G G T G G A T G T G C T G G C C T C G C G C G C 664

OY 227 L E U G I L U S P H E S E R A L A T H R A L A S E R S E R L E U G I Y P R O G I N L E U A R G L E U C Y S G I N 246
    |||||
DB 665 C T G A G A G A A T C T C A G C A C A T G C G G C A G T T C C T C G G C C C A C A G C C G C T C T G C A G 724

OY 247 V A L S E R G I L E U L A L A L E U A R P R O G I Y S E R S E R L E U A R G I L E A R G T H R L E U P R O T R P 266
    |||||
DB 725 G T G C T G G G T G T G G C C T G C G C A G G G T C T C C T G C G A T C C G A C C C T C C C C G G 784

OY 267 A L A H I S L E U Y S A L A A L A P R O P H E L E U T H R Y R P H E G L Y L E U P H E G I N V A L I S 284
    |||||
DB 785 G C C A T C T C A A G G C T G C C C T C T C A C C T A C T T C G A C T T C C A G G T T C A C 838
  
```

RESULT 4  
 AAA49717  
 ID AAA49717 standard; cDNA: 1353 BP.

```

XX AC AAA49717;
XX AC 25-SEP-2000 (first entry)
DT DT
XX DE Human PRO207 cDNA clone DNA30879-1152.
XX KW PRO207; human; antitumor; tumour; therapy; cytostatic;
KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;
KW uterine cancer; prostate cancer; lung cancer; bladder cancer;
KW central nervous system cancer; melanoma; leukemia; neoplasm; ss.
XX OS Homo sapiens.
XX FH Key
FH CDS Location/Qualifiers
FT CDS 58..807 /*cag- a
FT sig_peptide 58..177 /*cag- b
FT mat_peptide 178..804 /*tag- c
FT FT
XX PN WO200037638-A2.
XX PD 29-JUN-2000.
XX PF 02-DEC-1999; 99WO-US28565.
XX PR 22-DEC-1998; 98US-0113296.
PR 08-MAR-1999; 99WO-US05028.
PR 21-APR-1999; 99US-0130232.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21347.
XX PA (GETH ) GENENTECH INC.
XX PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;
PI Napier MA, Pitti RM, Wood WI;
DR P-PSDB; AAY95338.
XX WPI: 2000-442668/38.
XX PT Novel composition to inhibit neoplastic cell growth or for treating
PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,
PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO509 or
PT PRO866
XX PS Claim 20; Fig 3; 172pp; English.
XX CC The present sequence is that of cDNA clone DNA30879-1152
CC (ATCC 209358) encoding human PRO207 (see AAY95338), which shows
CC homology to several members of the tumor necrosis factor family,
CC especially human lymphotoxin (23.4%). The cDNA was identified in a
CC fetal kidney cDNA library following identification of an expressed
CC sequence tag with homology to human Apo-2 ligand. A claimed method
CC for inhibiting the growth of a tumour cell comprises exposing the
CC tumour cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224,
CC PRO328, PRO301, PRO526, PRO362, PRO509 or PRO866 (see
CC AAY95337-49), their agonists or chimeric polypeptides incorporating
CC them. The tumour is especially a cancer selected from breast,
CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and
CC central nervous system cancer, melanoma and leukemia. Nucleic
CC acids encoding PRO179 etc. are used in the recombinant production
CC of the antitumour polypeptides.
XX SQ Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;
  
```

Alignment Scores: 3.96e-90 Length: 1353  
 Pred. No.:



Score: 1366.00 Matches: 268  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 94.60% Indels: 0  
 DB: 21 Gaps: 0

US-09-245-198A-4 (1-284) x AAA49717 (1-1353)

```

OY 17 ArgSerLeuGlySerArgApGlyAlaValAlaArgAlaAlaGlnProAlaProMet 36
DB 1 CGATCCCTCGGGTCCCGGGATGGGGGGGGGTGAGGACAGACAGCCCCCCCCCATG 60
OY 37 AlaAlaArgArgSerGlnArgArgArgGlyArgArgGlyGluProGlyThrAlaLeu 56
DB 61 GCCGCCCTCGAGCCAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 120
OY 57 ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaVal 76
DB 121 GTCCCGCTCGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGG 180
OY 77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuVal 96
DB 181 GTCACTTTGGGAGCGGGGCGATCGCTGCTCCGCCAGAGCCCTGCCAGAGAGAGCTGGTG 240
OY 97 AlaGluGluAspGluAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspPro 116
DB 241 GCAGAGGAGAGACAGAGACCCGTCGAGACTGAATCCCGACAGAGAGAGAGAGAGATCCT 300
OY 117 AlaProPheLeuAsnArgLeuValAlaArgProArgSerAlaProGlyGlyArgLysThr 136
DB 301 GCGCCTTCCTGAACCGACTAGTTCGGGCTCGCAGAGTGCACCTAAAGCGGGAGAAACA 360
OY 137 ArgAlaArgArgAlaAlaAlaAlaHisTyrGluValHisProArgProGlyGlnAspGly 156
DB 361 CGGGCTCGAAGAGCATGCGACCCCATTAATGAGTTCACAGACCTGGAGAGCGCA 420
OY 157 AlaGlnAlaGlyValAlaAspGlyThrValSerGlyTyrPGLuGluAlaArgLysSerSer 176
DB 421 GCGGAGGAGAGTGTGGAGCGGAGCTGAGTGGCTGGAGAGAGCCAGATCAACACTTC 480
OY 177 SerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyr 196
DB 481 AGCCCTCGCGCTACACCGCCAGATCGGGAGATTATAGTCACCGGGGCTGGGCTCTAC 540
OY 197 TyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu 216
DB 541 TACCTGTACTGTGAGTGCATTTGATGAGGGGAGGCGTGTACTGTAAGCTGGACTTG 600
OY 217 LeuValAlaAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAspSer 236
DB 601 CTGGGTGGATGGTGTGCTGCTGGCCCTGGCGCTCCCTGGAGAGAAATTCACGCCCTCGGCGAGCT 660
OY 237 SerLeuGlyProGlnLeuArgLeuGlyGlnValSerGlyLeuLeuAlaLeuArgProGly 256
DB 661 TCCCTCGGGGCCCGACCTCGGCTCGCCAGGTGTGGGCTGTGGCCCTGGGCGAGGG 720
OY 257 SerSerLeuArgGlyLeuArgThrLeuProTrrAlaHisLeuLysAlaAlaProPheLeuThr 276
DB 721 TCTCTCCGTGGGATCCGACCTCCCTGGGCGCATTCACAGCTGCCCCCTTCTCCAC 780
OY 277 TyrPheGlyLeuPheGlnValHis 284
DB 781 TACTTCGAGACTTTCAGAGTTTAC 804

```

RESULT 5  
 ID ABR40255 standard; cDNA; 1353 BP.  
 XX ABR40255;

DT 15-JUN-2002 (first entry)  
 XX cDNA encoding human PRO207 polypeptide.

XX Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;  
 KW Leukaemia; neuronal disorder; stromal disorder; blastocoele disorder;  
 KW inflammatory disorder; immune disorder; angiogenic disorder;  
 KW gene therapy; cyostatic; neuroprotective; gene; ss.

XX Homo sapiens.

XX WO200153486-A1.

XX 26-JUL-2001.

XX 11-FEB-2000; 2000WO-US03565.

XX 08-MAR-1999; 99WO-US05028.

XX 11-MAR-1999; 99US-123972P.

XX 11-MAY-1999; 99US-13459P.

XX 02-JUN-1999; 99WO-US12252.

XX 22-JUN-1999; 99US-140650P.

XX 22-JUN-1999; 99US-140653P.

XX 20-JUL-1999; 99US-144758P.

XX 26-JUL-1999; 99US-145698P.

XX 28-JUL-1999; 99US-146222P.

XX 17-AUG-1999; 99US-149395P.

XX 31-AUG-1999; 99US-151689P.

XX 01-SEP-1999; 99WO-US20111.

XX 15-SEP-1999; 99WO-US21090.

XX 30-NOV-1999; 99WO-US28313.

XX 01-DEC-1999; 99WO-US28301.

XX 01-DEC-1999; 99WO-US28301.

XX 05-JAN-2000; 2000WO-US00219.

XX (GETH ) GENENTECH INC.

XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;

XX Marsters SA, Pan J, Pittl RM, Roy MA, Smith V, Stone DM;

XX Watanabe CK, Wood WI;

XX WPL 2002-205567/26.

XX P-PSDB; AA086129.

XX Thirty five nucleic acids encoding PRO polypeptides, useful for

XX treating benign or malignant tumours, leukaemias and lymphoid

XX malignancies, inflammatory, angiogenic and immunologic disorders -

XX Claim 50; Fig 3; 302pp; English.

XX The present invention relates to the isolation of novel human PRO

XX polypeptides and the polynucleotide sequences encoding them. The

XX PRO polypeptides, agonists, antagonists or anti-PRO antibodies are

XX useful for treating benign or malignant tumours (e.g. renal, kidney,

XX bladder, breast, etc), leukaemias and lymphoid malignancies, other

XX CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,

XX CC macrophagal, stromal and blastocoele disorders, inflammatory, immune

XX CC and angiogenic disorders. The polynucleotide sequences are also

XX useful in gene therapy. ABR40254-ABR40288 encode for the human PRO

XX polypeptides of the invention.

XX Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;

XX Alignment Scores:

XX Pred. No.: 3,966-90 Length: 1353

XX Score: 1366.00 Matches: 268

XX Percent Similarity: 100.00% Conservative: 0

XX Best Local Similarity: 100.00% Mismatches: 0

XX Query Match: 94.60% Indels: 0

XX DB: 24 Gaps: 0

XX US-09-245-198A-4 (1-284) x ABR40255 (1-1353)

OY 17 ArgSerLeuGlySerArgApGlyAlaValAlaArgAlaAlaGlnProAlaProMet 36  
 DB 1 CGATCCCTCGGGTCCCGGGATGGGGGGGGGTGAGGACAGACAGCCCCCCCCCATG 60

QY 37 AAlaAlaArgArgSerGlnArgArgGlyArgArgGlyGluProGlyThrAlaLeuLeu 56  
 Db 61 GCCGCCGCTGGAGCCAGAGCGAGGGGGCCGGGGGGAGCCGGGACCGCCCTGCTG 120  
 QY 57 ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaVal 76  
 Db 121 GTCCCGCTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 180  
 QY 77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuVal 96  
 Db 181 GTCAAGTTGGGGAGCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 QY 97 AlaGlnGluAspGlnAspProSerGluLeuAsnProGlnThrGlnGluSerGlnAspPro 116  
 Db 241 GCAGAGAGAGCAGAGACCGCTCGAAGCTGAATCCAGACAGAAAGAAAGCAGAGATCCT 300  
 QY 117 AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProGlyArgGlyThr 136  
 Db 301 GCGCCCTTCTGAAACCGACTAGTTGCGGCTGCGAGAAAGTGCACCTAAAGCCGGAAACA 360  
 QY 137 ArgAlaArgArgAlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGly 156  
 Db 361 CCGGCTCGAAGAGGATCCGAGCCCATTAATTCATCCACACACTGGACAGAGACGA 420  
 QY 157 AlaGlnAlaGlyValAspGlyThrValSerGlyTyrPglGluAlaArgIleAsnSerSer 176  
 Db 421 GCGCAGAGCAGCTGAGAGCGGACAGTGAAGTGTGGGAGGAGAAAGATCAACAGCTCC 480  
 QY 177 SerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValIleProAlaGlyLeuTyr 196  
 Db 481 AGCCCTTGCGCTACAAACCGCAGATCGGGAGTTTATGACACCCGGGCTGGCTTAC 540  
 QY 197 TyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu 216  
 Db 541 TACTGTACAGTGCAGGTGCACTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600  
 QY 217 LeuValAspGlyValIleAlaAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSer 236  
 Db 601 CTGCTGATGGTGTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660  
 QY 237 SerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGly 256  
 Db 661 TCCCTCGGGGCGCCAGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
 QY 257 SerSerLeuArgIleArgTyrLeuProTyrAlaHisIleuLysAlaAlaProPheLeuThr 276  
 Db 721 TCCCTCGGGGAGTCCGACCCCTCCCTGGGCTCATCTCAAGGCTGGCCCTTCTCACC 780  
 QY 277 TyrPheGlyLeuPheGlnValHis 284  
 Db 781 TACTTCGAGCTCTTCACAGTTTCAC 804  
 RESULT 6  
 ID AAV47613 standard; cDNA; 1236 BP.  
 AC AAV47613:  
 XX 27-OCT-1998 (first entry)  
 DE TNF related endothelium proliferative agent gene.  
 KW ss; TNF; endothelium proliferative agent; TREPA; wound healing; cancer;  
 tissue grafting; vascularisation; apoptosis; autoimmune; birth control.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT 1..750  
 FT /\*tag- a  
 FT /product- "TREPA"  
 XX

PN WC0835061-A2.  
 XX 13-AUG-1998.  
 PD 13-AUG-1998.  
 XX 12-FEB-1998; 98KO-US02859.  
 PF 12-FEB-1998; 98KO-US02859.  
 PR 10-FEB-1998; 98US-0021706.  
 PR 12-FEB-1997; 97US-0798692.  
 XX (ABBO ) ABBOTT LAB.  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 XX WALLEY SR.  
 PI  
 XX WPI: 1998-447255/38.  
 DR P-PSDB; AAW29745.  
 XX  
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 treatment of autoimmune disease, tumours and inflammation  
 PS Claim 11; Page 123-4; 142pp; English.

CC The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).

XX SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

#### Alignment Scores:

Pred.	7.37e-83	Length:	1236
Score:	1265.00	Matches:	248
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.60%	Mismatches:	0
Query Match:	87.60%	Indels:	0
DB:	19	Gaps:	0

US-09-245-198A-4 (1-284) x AAV47613 (1-1236)

QY 36 MetaAlaAlaArgArgSerGlnArgArgGlyArgArgGlyGluProGlyThrAlaLeu 55  
 Db 1 ATGGCCGCCCGCTCGAGACAGAAAGGAGGGGCGGGGGGGAGCCGGGACCGCCCTG 60  
 QY 56 LeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAla 75  
 Db 61 CTGGTCCCGCTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 120  
 QY 76 ValValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuVal 95  
 Db 121 GTGCTCAGTTGGGGAGCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180  
 QY 96 ValAlaGluGluAspGlnAspProSerGluLeuAsnProGlnThrGlnGluSerGlnAsp 115  
 Db 181 GTGCGACAGAGAGCAGAGACCGCTCGAAGCTGAATCCAGACAGAAAGAAAGCAGAGAT 240  
 QY 116 ProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProGlyArgGlyThr 135  
 Db 241 CCTGCGCTTCTGAAACCGACTAGTTGCGGCTGCGAAGTGCACCTAAAGCCGGGAAA 300  
 QY 136 ThrArgAlaArgArgAlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAsp 155  
 Db 301 ACAGGGGCTCGAAGAGGAGTCCAGCCCATTAATTCATCCACAGACTGGGACAGGAC 360

OY 156 GYAlAGlnAlaGlyValAspGlyThrValSerGlyTyrPGluGluAlaArgIleAsnSer 175  
 DB 361 GGAGGCGAGGAGGTGTGACGGGACAGTGAAGTGGCTGGAGAGAACAGATCAACAC 420  
 OY 176 SerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeu 195  
 DB 421 TCCACCCCTCGCGCTACACCGCCAGATCGGGAGTTTATGATCACCAGGCGCTGGGCTC 480  
 OY 196 TyrTyrLeuTyrCysGlnValHisPheAspGluGlyValAlaValTyrLeuLysLeuAsp 215  
 DB 481 TACTACCTGTACTGTACAGTGCATTTGATGAGGGGAAGGCTGTACCTGAAGCTGGAC 540  
 OY 216 LeuLeuValAspGlyValIleuAlaLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAla 235  
 DB 541 TTGCTGGTGGATGGTGTCTGCTGGCCCTGGCGCTCTCGAGAGAAATTCACAGCCCTGGGCG 600  
 OY 236 SerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgPro 255  
 DB 601 AGTTCCTCGGGCCCGACCTCGGCTCGCCAGGTGTGGGCTGTGGCCCTGGCGGCA 660  
 OY 256 GlySerSerLeuArgIleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeu 275  
 DB 661 GGGTCTCCCTGGAGATCGGACCTCCCTGGGCCCATCTCAAGGCTGCCCTTCCTC 720  
 OY 276 ThrTyrPheGlyLeuPheGlnValHis 284  
 DB 721 ACCTACTTCGGACTCTTCAGGTTTCAC 747

## RESULT 7

AAD04350 ID AAD04350 standard; cDNA; 1236 BP.

AC AAD04350;

DT 04-JUL-2001 (first entry)

DE Human TREPA (TNF related endothelium proliferative agent) cDNA.

KW Human: tumour necrosis factor; TNF; angiogenesis; wound healing;

KW TREPA: TNF related endothelium proliferative agent; tumour; metastasis;

KW grafting; vulnerrary; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..750 /tag- a /product- "Human TREPA (TNF related endothelium proliferative agent)"

PN US6207642-B1.

PD 27-MAR-2001.

PE 26-JUN-1998; 98US-0105343.

PR 12-FEB-1997; 97US-0798692.

PR 10-FEB-1998; 98US-0021706.

PA (ABBO ) ABBOTT LAB.

PI Wiley SR;

DR WPI: 2001-280760/29.

DR P-PSDB: AAE00891.

XX Inducing angiogenesis in mammal at desired sites for promoting wound

XX healing, by administering soluble fragment of extracellular domain of

XX tumor necrosis factor related endothelium proliferative agent protein

XX

CC The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts.  
 CC The present sequence is a cDNA clone ID #690050 encoding human TREPA.  
 XX

SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other:

## Alignment Scores:

Pred. No.:	7,37e-83	Length:	1236
Score:	1265.00	Matches:	248
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.60%	Mismatches:	0
Query Match:	87.60%	Indels:	0
DB:	22	Gaps:	0

US-09-245-198a-4 (1-284) x AAD04350 (1-1236)

OY 36 MetAlaAlaArgArgSerGlnArgArgGlyArgGlyGluProGlyThrAlaLeu 55

DB 1 ATGGCCGCCCGTCGGAGCCAGAACCGGAGGGGCCCGGGGGAGCCGGGACCCGCCCTG 60

OY 56 LeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAla 75

DB 61 CTGGTCCCGCTGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 120

OY 76 ValValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeu 95

DB 121 GTGGTCAGTTGGGAGCGGGGCGGCGATCGTCCGCCAGAGACCTGCCAGAGAGAGCTG 180

OY 96 ValAlaGluGluAspGlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAsp 115

DB 181 GTGGCAGAGAGAGACAGACACCGCTCGGAACGATCCCGACAGAGAAAGCCAGAT 240

OY 116 ProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLys 135

DB 241 CCGTGGCCCTTCTTCTACCGACTAGTTGGCTCGCAAGATGACCTTAAGAGCGGANA 300

OY 136 ThrArgAlaArgArgAlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAsp 155

DB 301 ACAGGGGCTCGAAGAGCGATGCGACCCATTTATGAAGTTCAATCCAGACCTGGACAGGAC 360

OY 156 GYAlAGlnAlaGlyValAspGlyThrValSerGlyTyrPGluGluAlaArgIleAsnSer 175

DB 361 GGAGGCGAGGCGAGGTGTGACGAGACAGTGAAGTGGCTGGAGAGCCAGAAATCAACAGC 420

OY 176 SerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeu 195

DB 421 TCCAGCCCTCGCGCTACACCGCCAGATCGGGAGTTTATGATCACCAGGCTGGGCTC 480

OY 196 TyrTyrLeuTyrCysGlnValHisPheAspGluGlyValAlaValTyrLeuLysLeuAsp 215

DB 481 TACTACCTGTACTGTACAGTGCATTTGATGAGGGGAAGGCGTGTACTGAGCTGGAC 540

OY 216 LeuLeuValAspGlyValIleuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAla 235

DB 541 TTGCTGGTGGATGGTGTCTGCTGGCCCTGGCGCTCGGAGAAATTCACAGCCACTGGGCG 600

OY 236 SerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgPro 255

DB 601 AGTTCCTCGGGCCCGACCTCGGCTCGCCAGGTGTGTGGCTGTGGCCCTGGGCGCA 660

OY 256 GlySerSerLeuArgIleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeu 275

DB 661 GGGTCTCCCTGGAGATCGGACCTCCCTGGGCCCATCTCAAGGCTGCCCTTCCTC 720

OY 276 ThrTyrPheGlyLeuPheGlnValHis 284

DB 721 ACCTACTTCGGACTCTTCAGGTTTCAC 747

RESULT 8  
 AAS03964  
 ID AAS03964 standard; DNA; 898 BP.  
 XX  
 AC AAS03964;  
 XX  
 DT 26-SEP-2001 (first entry)  
 XX  
 DE Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA.  
 XX  
 KM TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;  
 KM ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 KM retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 KM rubecosis; uveitis; macular degeneration; arthritis; rheumatism; ds;  
 KM corneal graft neovascularisation; psoriasis; metastatic condition;  
 KM malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
 KM preneoplastic condition; myocardial angiogenesis; wound granulation;  
 KM scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
 KM atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 KM peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;  
 KM fusion protein.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 FH Location/Qualifiers  
 FT 52..873  
 FT CDS /tag- a  
 FT /product- "Fusion protein comprising a growth hormone  
 FT leader, a leucine zipper multimerisation  
 FT domain, and human TWEAK extracellular  
 FT domain"  
 XX  
 PN WO200145730-A2.  
 XX  
 PD 28-JUN-2001.  
 XX  
 PE 19-DEC-2000; 2000MO-US34755.  
 XX  
 PR 20-DEC-1999; 99US-0172878.  
 PR 10-MAY-2000; 2000US-0203347.  
 XX  
 PA (IMM) IMMUNEX CORP.  
 XX  
 PI Walley SR.  
 DR WPI; 2001-417975/44.  
 DR P-PSDB; AAO03499.  
 XX  
 PI Modulating angiogenesis in a mammal for treating diseases mediated by  
 PI angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  
 PI peripheral tissue, by administering antagonist or agonist of TWEAK  
 PI receptor  
 XX  
 PS Example 1; Page 39-40; 46pp; English.  
 XX  
 CC The sequence represents a DNA from the expression vector  
 CC pDC409-LZ-TWEAK, which encodes a fusion protein comprising a growth  
 CC hormone leader, a leucine zipper multimerisation domain, and the  
 CC extracellular domain of human TWEAK. The fusion protein was used in  
 CC the isolation of human TWEAK receptor (TWEAKR)-expressing clones  
 CC from a COS cell human cDNA library. The TWEAK protein is a  
 CC member of the tumour necrosis factor (TNF) family and induces  
 CC angiogenesis. TWEAKR may therefore be used to screen for and develop  
 CC TWEAKR agonists and antagonists for the modulation of angiogenesis, to be  
 CC used in the treatment and diagnosis of human disease. The disorders  
 CC mediated by angiogenesis include ocular disorders characterised by ocular  
 CC neovascularisation such as diabetic retinopathy, neovascular glaucoma,  
 CC retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,  
 CC rubecosis, uveitis, macular degeneration and corneal graft  
 CC neovascularisation, and inflammatory diseases such as arthritis,  
 CC rheumatism and psoriasis. Other treatable diseases include malignant and  
 CC metastatic conditions such as sarcomas and carcinomas, benign tumours and

CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,  
 CC scleroderma, vascular adhesions, atherosclerotic plaque  
 CC neovascularisation, telangiectasia, wound granulation, coronary  
 CC atherosclerosis, peripheral atherosclerosis and ischaemia.  
 XX  
 SQ Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 1,32e-68 Length: 898  
 Score: 1065.00 Matches: 207  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 73.82% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-245-198a-4 (1-284) x AAS03964 (1-898)  
 QY 78 SerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuValAla 97  
 DB 250 AGTTTGGGAGCCGGGCAATCGCTGCCGCCAGAGACCTGCCAGAGAGAGCTGTGGCA 309  
 QY 98 GlnGluAspGlnAspProSerGluLeuAsnProGlnThrGlnGluSerGlnAspProAla 117  
 DB 310 GAGGAGAGACCGAGACCGCTGGAACTGAATCCCAACAGAAAGAACCCAGGATCTGCG 369  
 QY 118 ProPheLeuAsnArgLeuValArgProArgSerAlaProLysGlyArgLysThrArg 137  
 DB 370 CTTTCTGACCGACTAGTGTGGCCCTCCAGAAAGACACCTAAAGCCGGAAACACGG 429  
 QY 138 AlaArgArgAlaIleAlaAlaHisTYrGluValHisProArgProGlyLysAspGlyAla 157  
 DB 430 GCTCGAAGAGCGATCCAGCCCATTTATGAATTCATCCAGACCTGGACAGGAGCGAGCG 489  
 QY 158 GlnAlaGlyValAspGlyThrValSerGlyTTPGlnGluAlaArgIleAsnSerSerSer 177  
 DB 490 CAGGACAGGTGGACGAGGAGACGAGTGGCTGGAGAGAACCCAGATCAACAGCTCCAGC 549  
 QY 178 ProLeuArgTYrAsnArgGlnIleGlyLysPheIleValThrArgAlaGlyLeuTYr 197  
 DB 550 CCTCGCGCTACACACCGCCAGATCGGGAGATTATAGTCACCGGGGCTGCTACTAC 609  
 QY 198 LeuTYrCYsGlnValHisPheAspGluGlyValAlaTYrLeuLysLeuAspLeu 217  
 DB 610 CTGTACTGTAGGTCATTTGATGAGGGAGAGGCTGTACTCAAGCTGAGCTGTG 669  
 QY 218 ValAspGlyValLeuAlaLeuArgCysLeuGlnGluPheSerAlaThrAlaAspSer 237  
 DB 670 GTGATGCTGTGCTGGCCCTGCGCTGCTGGAGGATTCACGCCACCTGGGCAATTCC 729  
 QY 238 LeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySer 257  
 DB 730 CTCGGGCCCCAGCTCCGCTGCTGCGAGTGTGGGCTGTGGCCCTGGGGCCAGGCTCC 789  
 QY 258 SerLeuArgIleArgThrLeuProTTPAlaHisLeuLysAlaAlaProPheLeuThrTYr 277  
 DB 790 TCCCTCGGATCCGACACCTCCCTGGGCCCATCTCAAGGCTGCCCTTCTCACTAC 849  
 QY 278 PheGlyLeuPheGlnValHis 284  
 DB 850 TTGGACTCTTCCAGGTTCCAC 870  
 RESULT 9  
 AAX23424  
 ID AAX23424 standard; DNA; 1030 BP.  
 XX  
 AC AAX23424;  
 XX  
 DT 18-JUN-1999 (first entry)  
 XX  
 DE Human TNRL3 DNA.  
 XX  
 KM Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;  
 KM developmental abnormality; gestational abnormality; prostate cancer;

KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW Apoptosis; human; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..627  
 FT CDS /tag= a  
 FT /product= "TNRL3"  
 XX  
 PN W09911791-A2.  
 XX  
 PD 11-MAR-1999.  
 XX  
 PF 04-SEP-1998; 98WO-US18393.  
 XX  
 PR 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNITV ) UNITV WASHINGTON.  
 XX  
 PI Chaudhary PM;  
 XX  
 DR WPI; 1999-205191/17.  
 DR P-PSDB; AAM93590.  
 XX  
 PT New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 XX  
 PS Example VII; Fig 13A; 156pp; English.  
 XX  
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC moiety. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the changer in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active  
 CC fragments and APO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
 CC activity. The method is performed in vivo or in vitro. APO polypeptides  
 CC are all useful as immunogens for preparing antibodies. APO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 XX  
 SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 3, 01e-68 Length: 1030  
 Score: 1062.00 Matches: 207  
 Percent Similarity: 99.52% Conservative: 0  
 Best Local Similarity: 99.52% Mismatches: 1  
 Query Match: 73.55% Indels: 0  
 DB: 20 Gaps: 0  
 US-09-245-198a-4 (1-284) x AAX23424 (1-1030)  
 QY 77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuVal 96  
 DB 1 GTGAGTTGGGAGCCGGGATCGCTGCGCCAGAGAGCTCCAGAGAGAGCTGCTG 60  
 QY 97 AlaGluGlnAspGlnAspProSerGluLeuAsnProGlnThrGlnGlnSerGlnAspPro 116  
 DB 61 GCAGAGGAGGAGCAGAGCCGCTCGAATGAAATCCAGAGAGAAAGCAGAGATCCT 120

QY 117 AlaProPheLeuAsnArgLeuValArgProArgSerAlaProIlysglyArglyThr 136  
 DB 121 GCGCTTTCTGAAACCGACTAGTTCGGCTCCGAGAAAGTCCACTAAAGCCGGAATA 180  
 QY 137 ArgAlaArgArgAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 156  
 DB 181 CGGCTCGAAGAGCAGATCCGAGCCATTATGAGTTATCATCCAGACTGACAGAGAG 240  
 QY 157 AlaGlnAlaGlyValAspGlyThrValSerGlyTTPGluGluAlaArgIleAsnSer 176  
 DB 241 GCCAGGACAGCTGTCAGCGGACGAGCAGTGGCTGGAGGACCAAGAAACACACCTCC 300  
 QY 177 SerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyr 196  
 DB 301 AGCCCTCTCGCTACTACCGCCAGATCGGAGATTATATCTACCCGCGCTGCTTAC 360  
 QY 197 TyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuYsLeuAspLeu 216  
 DB 361 TACCTGACTGTGAGGTCGACTTGTATGAGGAGGAGGCTGTACCTGAAAGCTGACTTG 420  
 QY 217 LeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleSer 236  
 DB 421 CTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480  
 QY 237 SerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGly 256  
 DB 481 TCCCTGGGCGCCAGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540  
 QY 257 SerSerLeuArgIleArgThrLeuProThrAlaHisLeuYsAlaAlaProPheLeuThr 276  
 DB 541 TCCCTCCCTGGGAGATCCGCAACCTCCCTGGGAGATCCCTGGGAGATCCCTGGGAG 600  
 QY 277 TyrPheGlyLeuPheGlnValHis 284  
 DB 601 TACTCGAGCTCTCCAGGTTTCAC 624  
 RESULT 10  
 ID AAV18599 standard; cDNA: 1168 BP.  
 XX  
 AC AAV18599;  
 XX  
 DT 21-JUL-1998 (first entry)  
 XX  
 DE Mus musculus tumour necrosis factor related ligand (TNRL) gene.  
 XX  
 KW TNRL; tumour necrosis factor related ligand; tnfr; treatment;  
 KW cancer; autoimmune disease; immune system; stimulation; suppression;  
 KW graft rejection; ds.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT 2..679  
 FT CDS /tag= a  
 FT /note= "tumour necrosis factor related ligand"  
 XX  
 PN W09805783-A1.  
 XX  
 PD 12-FEB-1998.  
 XX  
 PF 07-AUG-1997; 97WO-US13945.  
 XX  
 PR 18-MAR-1997; 97US-0040820.  
 PR 07-AUG-1996; 96US-0023541.  
 PR 18-OCT-1996; 96US-0028515.  
 XX  
 PA (BIOI ) BIOGEN INC.  
 PA (UYGE-) UNITV GENEVA FACULTY MEDICINE.  
 XX  
 PI Browning JL, Chicheportliche Y;  
 XX

DR WPI; 1998-145619/13.  
 DR P-PSDB; AAM47524.  
 XX Tumour necrosis factor related ligand - useful for, e.g. treating  
 PT cancer, auto-immune disease and immune responses to tissue grafts  
 PS  
 XX Claim 2; Pages 45-46; 69pp; English.

CC The sequence is that encoding mouse tumour necrosis factor related  
 CC ligand (TRELL). TRELL or active fragments can be included with a  
 CC carrier in pharmaceutical compositions to treat cancer, auto-immune  
 CC diseases or immune responses to tissue grafts, or to stimulate or  
 CC suppress the immune system. It is useful to screen for TRELL  
 CC receptors, by labelling with a detectable label and screening  
 CC compositions for binding. Agents interfering with TRELL-receptor  
 CC binding can also be screened for, can then be administered,  
 CC optionally with interferon- gamma, to induce cell death or  
 CC treat, suppress or alter immune responses (especially involving human  
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
 CC receptor. The DNA sequence can be used in gene therapy for  
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
 CC autoimmune and inflammatory diseases or inherited genetic disorders,  
 CC by introducing into cells, and expressing, therapeutically effective  
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
 CC It may also be of use in the preparation of probe probes for  
 CC screening natural/synthetic DNAs for TRELL-encoding sequences  
 CC and for antisense therapy.

SO Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other;

Alignment Scores:  
 Pred. No.: 3.8e-65 Length: 1168  
 Score: 1020.00 Matches: 199  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 88.84% Mismatches: 16  
 Query Match: 70.64% Indels: 0  
 DB: Gaps: 0

US-09-245-198A-4 (1-284) x AAV18599 (1-1168)

QY 61 LeuGlyLeuGlyLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaValSerLeuGly 80  
 DB 5 CTGACCTGGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGG 64  
 QY 81 SerArgAlaSerLeuSerAlaGlnGlnProAlaGlnGlnGlnGlnGlnGlnGln 100  
 DB 65 AGCTGGGCAAGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 124  
 QY 101 GlnAspProSerGlnLeuAsnProGlnThrGlnGlnGlnGlnGlnGlnGln 120  
 DB 125 CGGAGCGCCCTGAACCTGAATCCCAAGACAGAGAAAGCCAGATGCTGCTGCTG 184  
 QY 121 AsnArgLeuValArgProArgArgSerAlaProGlyArgGlyArgAlaArgArg 140  
 DB 185 GAAACACTAGTCCGCGCTCGAAGAGTCTCTTAAGCCGGAAGCGCGCTGCGCGA 244  
 QY 141 AlaIleAlaIleHisTyrGlnValHisProArgProGlyGlnAspGlyAlaGln 160  
 DB 245 GCTATTGACGCCATTAAGAGGTTCACTCGGCGAGAGAGAGAGAGAGAGAGAG 304  
 QY 161 ValAspGlyThrValSerGlyTyrGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 180  
 DB 305 GTGGATGGACAGTACGAGCGCTGGAGAGACCAAAATCAACAGCTCCAGCCCTGCGC 364  
 QY 181 TyrAsnArgGlnIleGlyGlnPheIleValThrArgAlaGlyLeuTyrTyrLeuTyr 200  
 DB 365 TACACCGCCAGATTTGGGATTTACAGTCATCAGGGGTGGCTCTACTACTACT 424  
 QY 201 GlnValHisPheAspGlnGlyGlyAlaValTyrLeuTyrLeuAspLeuValAspGly 220  
 DB 425 CAGGTGACCTTTGATGAGGAGAAAGCTGTCTACCTGAAGCTGAGCTGCTGTAACGGT 484  
 QY 221 ValLeuAlaLeuArgCysLeuGlnGlnPheSerIleThrAlaIleSerLeuGlyPro 240

DB 485 GTGCTGGCCCTGGCTGGCTGGCAAGATTCAGCCACAGACAGCAAGCTCTCTGGGCC 544  
 QY 241 GlnLeuArgLeuGlyGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArg 260  
 DB 545 CAGCTCCGTTTGTGTCAGAGGTGTGGGCTGTCCCTGGCGGAGGCTTCTCCCTGG 604  
 QY 261 IleArgThrLeuProTrrPalaHisLeuValAlaProPheLeuThrTyrPheGlyLeu 280  
 DB 605 ATCCGACACCTCCCTCCGAGGCTCATCTTAAGGCTGCCCTCTTCACTTCTTGA 664  
 QY 281 PheGlnValHis 284  
 DB 665 TTTCAAGTTCAC 676  
 RESULT 11  
 AAX23425  
 ID AAX23425 standard; DNA; 701 BP.  
 XX AAX23425;  
 AC 18-JUN-1999 (first entry)  
 DT  
 XX Mouse TNRL3 DNA.  
 DE  
 KW Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;  
 KW developmental abnormality; gestational abnormality; prostate cancer;  
 KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW apoptosis; mouse; ss.  
 XX  
 XX Mus sp.  
 OS  
 XX Key Location/Qualifiers  
 FH 1..636  
 FT CDS  
 FT /tag- a  
 FT /product- "TNRL3"  
 FT  
 XX W09911791-A2.  
 PN  
 XX 11-MAR-1999.  
 PD  
 XX  
 XX 04-SEP-1998; 98WO-US18393.  
 PE  
 XX 05-SEP-1997; 97US-0924634.  
 PR  
 XX (UNIM ) UNIV WASHINGTON.  
 PA  
 XX Chaudhary PM;  
 PL  
 XX WPI; 1999-205191/17.  
 DR P-PSDB; AAM93591.  
 DR  
 XX New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 PT  
 XX Example VII; Fig 13B; 156pp; English.  
 PS  
 XX This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC moiety. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the change in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active

fragment and Ap04 signal transducer molecules that specifically interact with a cytoplasmic domain of Ap04 and detecting a change in level of Ap04 activity. The method is performed in vivo or in vitro. Apo polypeptides are all useful as immunogens for preparing antibodies. Ap04 is also useful for diagnosis/treatment of developmental or gestational abnormalities. Ap08 was transfected to human breast carcinoma cell line MCF-7, and induced apoptosis.

**SQ** Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;

### Alignment Scores:

Pred. No.:	1,256-61	Length:	701
Score:	968.00	Matches:	188
Percent Similarity:	93.36%	Conservative:	9
Best Local Similarity:	89.10%	Mismatches:	14
Query Match:	67.04%	Indels:	0
DB:	20	Gaps:	0

US-09-245-198A-4 (1-284) x AAX23425 (1-701)

QY	74	LeuAlaValValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGlu	93
Db	1	CTGGTCGGGGTCACCTCGGGAGCTGGGCAACGCTGCTCCACGAGGACCTTTCCAGAG	60
QY	94	GluLeuValAlaGlnGluGlnAspGlnAspProSerGluLeuAsnProGlnTrpGluGluSer	113
Db	61	GAGCTGACAGCAGAGAGAACCGCGGGAGACCCCTGAACTGAATCCCGACAGAGAAAGC	120
QY	114	GlnAspProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysGly	133
Db	121	CAGGATGGTACCTTTCTTGGAACAACATGATCGGGCTCGAAGAAAGTGCTCTAAAGG	180
QY	134	ArgLysThrArgAlaArgArgAlaAlaAlaAlaHisTrpTrpValHisProArgProGly	153
Db	181	CGGAAGGCGCGGCTCGCGAGCTATTGCACGCCATTATGAGTCACTCTCGGCCAGGA	240
QY	154	GlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTrpGlnGlnAlaArgile	173
Db	241	CAGGATGGAGCACAAGCAGGTGTGATGGGACAGTAGTGGCTGGGAGAGACCAAAATC	300
QY	174	AsnSerSerProLeuAlaGlyTrpAsnArgGlnIleGlyGluPheIleValThrArgAla	193
Db	301	AACAGCTCCACGCTCTGCGCTACGACCGCCAGATGGGAAATTTACAGTCACAGGGCT	360
QY	194	GlyLeuTrpTrpLeuTrpCysGlnValHisPheAspGlnGlyLysAlaValTrpLeuLys	213
Db	361	GGGCTTACTACTCGTACTGTGACAGGTGCATTTGATGAGGAAAGGCTGTACCTGAAG	420
QY	214	LeuAspLeuLeuValAspGlyValLeuAlaLeuArgCysLeuGlnPheSerAlaTrp	233
Db	421	CTGGACCTGGTGGTGAACGGGTGCTGCGCTGCGCTGCGAAGAAATTTCTAAGCCACA	480
QY	234	AlaAlaSerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeu	253
Db	481	GCACCAAGCTCTCTCGGGCCCCACAGCTCCGTTTGCCAGGATGCTGAGGCTGTGGCGCTG	540
QY	254	ArgProGlySerSerLeuArgIleArgThrLeuProTrpAlaHisLeuLysAlaAlaPro	273
Db	541	CGGGCAGGGGTCTTCCCTTGGATTCGCAACCTCCCTGGGGCTCATCTTAAGGTGCCCC	600
QY	274	PheLeuThrTrpPheGlyLeuPheGlnValHis	284
Db	601	TTCCTAACCTACTTGGACGCTTTCACAAAGTTCAC	633
RESULT 12			
AAV24140			
ID AAV24140 standard; DNA; 2856 BP.			
XX AAV24140;			
XX			
DT 28-SEP-1998 (first entry)			
XX Homo sapiens BARD1 exon 1.			
DE			

XX			BARD1; ring protein; BRCA1; breast cancer; risk; diagnosis; exon; ss.
KX			Homo sapiens.
XW			
OS			
XX			
FH	Key	Location/Qualifiers	
FT	exon	2031..2188	
FT		//tag=a	
PT		/number=1	
XX			
PV			
PN	WO9812327-A2.		
XX			
PD	26-MAR-1998.		
XX			
PF	19-SEP-1997;	97WO-US16842.	
XX			
PR	04-APR-1997;	97US-0042985.	
PR	20-SEP-1996;	96US-0025296.	
PR	03-APR-1997;	97US-0042611.	
XX	(TEKA ) UNIV TEXAS SYSTEM.		
PA			
XX	Baer R., Bowcock AM;		
PI			
DR	WPI; 1998-230317//20.		
XX	DNA sequence encoding BARD1, B123, BE2, BE14, BE31 or BE445 - which is breast cancer antigen, BRCA1, binding proteins are useful to identify patient having or at risk of developing cancer		
PS	Claim 25; Page 307-308; 348pp; English.		

The sequence is that of containing a BARD1 exon which can be used in the preparation of the recombinant breast cancer antigen, BRCA1, binding proteins BARD1, B123, BE2, BE14, BE31 or BE445, or a composition for the detection of a BARD1, B123, BE2, BE14, BE31 or BE445 nucleic acid sequence, specifically a wild type BARD1 composition for the detection or purification of BRCA1, useful to identify a patient having, or at risk of developing cancer. BARD1 can be used in the preparation of an anti-BARD1 antibody, and in the detection and purification of a BRCA1 protein. BARD1, B123, BE2, BE14, BE31 or BE445 can be used in the identification of a binding protein agonist or antagonist that alters the binding of BARD1, B123, BE2, BE14, BE31 or BE445 to BRCA1 or the biological activity of the BRCA1-BARD1, B123, BE2, BE14, BE31 or BE445 complex. The antibodies can be used to detect BARD1, B123, BE2, BE14, BE31 or BE445, a specific anti-BARD1 antibody can be used to identify a patient having or at risk of developing cancer.

Sequence 2856 BP; 616 A; 778 C; 786 G; 674 T; 2 other;

### Alignment Scores:

Pred. No.:	4.74	Length:	285
Score:	127.00	Matches:	84
Percent Similarity:	36.00%	Conservative:	33
Best Local Similarity:	25.85%	Mismatches:	122
Query Match:	8.80%	Indels:	86
DB:	19	Gaps:	13

US-09-245-198A-4 (1-284) x AAV24140 (1-2856)

OY	14	ProLeuProArgSerLeuGlySerArgAspGlyGlyAlaVal	-----	27
			:::	
Db	1788	CCCACTGGCCAGCGCGCTCCGGGGCCCGCCCTGTGGCTCGAGTCCATT		1844
OY	28	-----ArgGlnAlaGlnPro-----	-----ProAlaProMetAlaAla	38
		:	:     :	
Db	1848	TGGAGCATGCGGCCCGCCCGCCGCCCTCCCGCGCACCCCGCCCTCTGGCGGC		1907
OY	39	ArgArgSerGlnArgArgArgGlyArgArgGlyGluProGly	ThAlaLeuLeuValPro	58
		:     :		
Db	1908	CGCCGTCCTCCAGACCCCGGGAAAGACTTGGCCGGTTTGGAGTCGCTGCCTGCACGCTTCCCT		1967
OY	59	LeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValSer		78





OY 131 ProlysglyArgLysThrArgAlaArgAlaIleAlaHisIleValHisPro 150  
 DB 742 -----GGCTGGGGGGCCAGACATCGAGCG-----CCT 771  
 OY 151 ArgProGlyGln-----AspGlyAlaGln-----158  
 DB 772 GCGTCTGTGGCGCGGAGCGTCCCGGGGTTCTGGAGCGCGTGGCAACTTCGGGTC 831  
 OY 159 -----AlaGlyValAspGlyThrValSer 166  
 DB 832 CGGAGCGACATCTCTACGTGCCAGCAGCAGCGCCCTGCGCCAGCGAGCATGTGAG 891  
 OY 167 -GlyTrpGluGluAlaArgIleAsnSerSerProLeuArgThrValSerArgGlnIleG 186  
 DB 892 GAGCTTGAGAGCGCCATCTCGAATCTCCGTCGCCCGCGTTCGGCGCCCTGCAGGC 951  
 OY 186 yGluPheIleValThrArgAlaGlyLeuTyrrTyrrLeuTyrrGlnValHisPheAspG 206  
 DB 952 GCCGCTGTGTGTCACGAGGAGCGGAGCTAC-----982  
 OY 206 uGlyLysAlaValTyrrLeuLysLeuAspLeuValAspGlyValLeuAlaLeuArgCy 226  
 DB 983 -----TCCGCTGCGCGAGCAGCATGAGCGCTTCTGCGCGCTCGGCTCGCGAGGC 1038  
 OY 226 sLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGly-----ProGlnLe 242  
 DB 1039 CGCGGAGCCCTTCAGAGAGCGGCGCTCCCAAGATGGCGAGCTTCGCGCGCGATGCT 1098  
 OY 242 uArgLeuGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleAr 262  
 DB 1099 GCCCATGACGTACATCAGCGGGGTC-----GGCAACAGAGCGCGCTGCG 1143  
 OY 262 gThrLeuProThrAlaHis---LeuLysAlaAlaProPheLeuThrTyrrPheGly 279  
 DB 1144 GCTGTGTCTGTGGGCGAAGACTTCGACGCGCGCGAGGTGGCGCGGCGG 1198

RESULT 14  
 AAV69289  
 ID AAV69289 standard; cDNA; 2832 BP.  
 XX  
 AC AAV69289;  
 DT 01-MAR-1999 (first entry)  
 XX  
 DE Murine JNK-interacting protein 1 (JIP-1) cDNA.  
 XX  
 KW JIP-1; JNK-interacting protein; c-Jun NH2-terminal kinase;  
 KW signal transduction; inhibitor; mouse; neurodegenerative disease;  
 KW Parkinson's disease; Alzheimer's disease; blood clot; stroke;  
 KW malignancy; cancer; leukemia; autoimmune disease; inflammation;  
 KW apoptosis; therapy; diagnosis; ds.  
 XX  
 OS Mus sp.  
 XX  
 FH Key location/Qualifiers  
 FT CDS 180..2162  
 FT /\*\*tag= a  
 XX  
 XX MO9849188-A1.  
 XX  
 PD 05-NOV-1998.  
 XX  
 PF 28-APR-1998; 98WO-US08513.  
 XX  
 PR 28-APR-1997; 97US-0819177.  
 XX  
 PA (UYMA-) UNIV MASSACHUSETTS.  
 XX  
 PI Davis RJ, Dickens M;  
 XX  
 DR WPI; 1999-024042/02.  
 DR P-PSDB; AAM81525.

XX  
 PT c-Jun NH2-terminal kinase (JNK)-interacting protein 1 - used to  
 PT treat neurodegenerative disease, blood clot, leukemia, autoimmune  
 PT disease, and inflammation  
 XX  
 PS Claim 11; Page 66-70; 95pp; English.  
 XX  
 CC This cDNA sequence codes for murine JNK-interacting protein 1  
 CC (JIP-1, see AAM81625), a novel cytoplasmic anchor protein that  
 CC specifically binds to and inhibits the biological effects of JNK  
 CC (c-Jun NH2-terminal kinase), including the initiation of apoptosis  
 CC and oncogenic transformation. To isolate the clone, a yeast  
 CC two-hybrid method was used to screen a mouse embryo cDNA library to  
 CC identify proteins that interact with JNK. 7 overlapping partial  
 CC clones were isolated, and the full-length cDNA was subsequently  
 CC obtained by screening a mouse brain lambda ZAPI cDNA library. The  
 CC invention provides JIP-1 nucleic acids and polypeptides, expression  
 CC vectors and host cells. The JIP-1 polypeptide and nucleic acids  
 CC (including antisense and ribozyme molecules) can be used in the  
 CC manufacture of a medicament for treating a pathological condition  
 CC associated with abnormal expression or activity of JNK, such as a  
 CC neurodegenerative disease (selected from Parkinson's disease and  
 CC Alzheimer's disease), a blood clot, stroke, malignancy, leukemia,  
 CC an autoimmune disease or inflammation (all claimed).  
 XX

SO Sequence 2832 BP; 569 A; 888 C; 840 G; 535 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 7.12 Length: 2832  
 Score: 124.50 Matches: 72  
 Percent Similarity: 36.06% Conservative: 25  
 Best Local Similarity: 26.77% Mismatches: 73  
 Query Match: 8.62% Indels: 100  
 DB: 20 Gaps: 12

US-09-245-198a-4 (1-284) x AAV69289 (1-2832)  
 OY 1 MetSerLeuLeuAspPheGluIleSerAlaArgLeuProLeuProArgSerLeuGly 20  
 DB 20 ATAGCTTGATATGCTGTCCGGAGCCCG-----CCCGGCCAGCTCAGTCCGAA 70  
 OY 21 SerArgAspGlyGlyAlaValArgGlnAla-----30  
 DB 71 CCCCGCGCGCGCGCGCGCTCCACCGCTCCGCTGTCGCGCGCGCGCGCGCGCG 130  
 OY 31 GlnProPheAlaPheMetAlaAlaArgArgSerGln-----ArgArgArgGlyArg 47  
 DB 131 CGCGCTCCCGCGCGCGCTCTCCGCGGATGCGCAGGGCTGCACCCCGGAATGGCGAGCG 190  
 OY 48 ArgGlyGluProGlyThrAlaLeuLeuValProLeu-Ala-----LeuGlyLe 63  
 DB 191 AGAGAGCGGCTGTGGCGGGGGCGCGCGCTCCCGACCGGCGCTCCCATTCCTGGACT 250  
 OY 63 u-----63  
 DB 251 GCACATCGCTGCGCTCCCAATTCAGGCTCACCCATGACATCAGCTGAGAGTTTGA 310  
 OY 64 -----GlyLeuAlaLeuAlaCys-----69  
 DB 311 GGATGAAGACCTTTCGAGATCACTGACGAGTGTGATCAGCTGATCAAGCAAGACAC 370  
 OY 70 -LeuGlyLeu-----LeuLeuAlaValValSerLeuGlySerArgAl 83  
 DB 371 CTTGTCTCTCGGGCGCGCGCGCGCGGCTGTGTCTGCGGGTACGACGCGAGCGGGG 430  
 OY 83 aSerLeuSerAlaGlnGluProAlaGlnGluLeuVal-----96  
 DB 431 GAGCGCGCTGCAGCGGAGATGCTGCAGATGACCTGTGTCAGCGGAGTGCACATTC 490  
 OY 97 -----AlaGluGluAspGlnAsp-----ProSerG 105  
 DB 491 GGGCGCCGAGGAGCAGGAGGAGGAGAGAGAGAGAGTCTGCTGCCAAGCAGCAGGAGT 550



```

Oy 224 uArgCysLeuGIuGIuPheSerAlaThrAlaAlaSerSerLeuGIuProGIuLeuArgLe 244
    |||
    -----TTCGTGCGCGGGGGGCTACG-----GGACCC----- 31378
Db 31408 CCGG-----

Oy 244 uCysGIuValSerGIuLeuAlaLeuArgProGIuSerSerLeuArgIleArgThrLe 264
    |||
    -----GGCCTTGTCGCTGAACGCGCTGCTTTGTTCACCGGCTCACCCACCC 31331
Db 31377 -----

Oy 264 uProTTPalaHisLeuLysAlaAlaPro 273
    |||
    -----:::||| 31330
Db 31330 CCCCCCACTCATACACGCGGTGCCCC 31303

```

Search completed: May 8, 2003, 03:43:17  
 Job time : 248.35 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 8, 2003, 02:22:12 ; Search time 2227.92 Seconds

(without alignments)  
3709.824 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 1444  
Sequence: 1 MSLLDFEISARRLPRLRSIG.....PNAHLKAPPLTYFGLEQVH 284

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ .p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool/US09245198/runat\_06052003\_170239\_9615/app\_query\_fasta\_1.846  
-DB=GenEmbl -OFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45  
-LOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=psio -NOR=ext -HEA=SIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USFR=US09245198.ecgn\_1\_1\_3073.etrnat\_06052003\_170239\_9615 -NCPV=6 -ICPV=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_iny:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rdd:\*  
36: em\_htg\_rdd:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1366	94.6	1353	6 AX201324	AX201324 Sequence
2	1366	94.6	1368	9 AF055872	AF055872 Homo sapi
3	1331.5	92.2	1651	9 BC019047	BC019047 Homo sapi
4	1298	89.9	1306	9 AF030099	AF030099 Homo sapi
5	1265	87.6	1236	6 AR140407	AR140407 Sequence
6	1066	73.8	898	6 AX180714	AX180714 Sequence
7	1020	70.6	1168	10 AF030100	AF030100 Mus muscu
8	589.5	40.8	177703	2 AC016876	AC016876 Homo sapi
9	589.5	40.8	215795	2 AC127470	AC127470 Pan trogl
10	557	38.6	177535	2 AC130192	AC130192 Sus scrof
11	555.5	38.5	161428	2 AC126925	AC126925 Canis fam
12	555.5	38.5	203083	2 AC069459	AC069459 Mus muscu
13	555.5	38.5	234182	10 AL603707	AL603707 Mouse DNA
14	495	34.3	153553	2 AC126921	AC126921 Bos tauru
15	489.5	33.9	138792	2 AC119115	AC119115 Rattus no
16	385.5	26.7	184026	2 AC098923	AC098923 Rattus no
17	214	14.8	203281	2 AC126237	AC126237 Canis fam
18	145.5	10.1	187835	8 AP002482	AP002482 Oryza sat
19	140	9.7	42210	1 SC1C2	AL031124 Streptomy
20	138.5	9.6	190050	1 AL646059	AL646059 Rattus no
21	136.5	9.5	242930	1 AC105470	AC105470 Rattus no
22	135	9.3	6397	1 AF002222	AF002222 Pseudomon
23	135	9.3	10728	1 AE004651	AE004651 Pseudomon
24	134	9.3	10965	1 AE004740	AE004740 Pseudomon
25	130	9.0	40352	1 SCC61A	AL356595 Streptomy
26	129	8.9	5751	1 SGGNASTRS	X89010 S. glaucosce
27	129	8.9	16185	1 SCC46	AL131964 Streptomy
28	129	8.9	25459	1 SGAJ5985	AT006585 Streptomy
29	129	8.9	32941	1 SCMI0	AL133469 Streptomy
30	129	8.9	43961	1 SCMI1	AL133422 Streptomy
31	128	8.9	137833	9 AC016708	AC016708 Homo sapi
32	127	8.8	194304	9 AC016708	AC016708 Homo sapi
33	126.5	8.8	207050	1 AL646053	AL646053 Rattus no
34	125.5	8.7	191221	2 AP005292	AP005292 Oryza sat
35	125.5	8.7	999	1 HSBARD11	AF038034 Homo sapi
36	125.5	8.7	28795	1 AVIN1FC	M20568 A. vinelandi
37	125.5	8.7	87144	8 AP003822	AP003822 Oryza sat
38	125	8.7	153595	2 AP004272	AP004272 Oryza sat
39	125	8.7	30590	1 SC4G2	AL031371 Streptomy
40	124.5	8.6	191221	2 AP005292	AP005292 Oryza sat
41	124.5	8.6	36734	1 SC6G10	AL049497 Streptomy
42	124.5	8.6	98504	2 AP003876	AP003876 Oryza sat
43	124.5	8.6	107132	2 AP003911	AP003911 Oryza sat
44	124.5	8.6	154746	14 HSY2HG52	286099 Herpes simp
45	124	8.6	176580	2 AP003634	AP003634 Oryza sat
			3957	6 A45258	A45258 Sequence 2

## ALIGNMENTS

RESULT 1

AX201324	LOCUS	AX201324	1353 bp	DNA	linear	PAT 30-AUG-2001
	DEFINITION	Sequence 3 from Patent WO0153486.				
	ACCESSION	AX201324				
	VERSION	AX201324.1	GI:15391154			
	KEYWORDS					
	SOURCE	human.				
	ORGANISM	Homo sapiens				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
REFERENCE		1 (bases 1 to 1353)				
AUTHORS		Ashkenazi,A.J., Goddard,A., Godowski,P.J., Gurney,A.L.,				
		Hillan,K.J., Marsters,S.A., Pan,J., Plii,R.M., Roy,M.A., Smith,V.,				
TITLE		Stone,D.M., Watanabe,C.K. and Wood,W.I.				
JOURNAL		Compositions and methods for the treatment of tumour				
		Patent: WO 0153486-A 3 26-JUL-2001;				
		Genentech, Inc. (US)				
FEATURES		location/Qualifiers				
source		1..1353				
		/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
BASE COUNT		257 a	443 c	389 g	264 t	
ORIGIN						

Alignment Scores:	
Pred. No.:	4.65e-63
Score:	1366.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	94.60%
OB:	6
Length:	1353
Matches:	268
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-245-198A-4 (1-284) x AX201324 (1-1353)

QY	17	ArgSerLeuGlySerArgAspGlyGlyValValaIarGlnAlaGlnProProAlaPromet	36
Db	1	CGATCCCTCGGGATCCCGGGATGGGGGGCGGTGAAGCCAGGACAGCCCCCGCCCATATG	60
QY	37	AlaAlaArgArgSerGlnArgArgArgGlyIarGArgGlyIupProGlyTThrAlaLeuLeu	56
Db	61	CGCGCCCTCGGAGCCAGAGCGGAGGGGGCGCCGGGGGGAGCCGGGCAACCGCCCTGCTTG	120
QY	57	ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaIacSLeuGlyLeuLeuLeuAlaVal	76
Db	121	GTCCCGCTCGGCGTGGGCGCTGGGCGCTGGCGCGTGGGCTGCTCGGCGCTCTGCGCGCTG	180
QY	77	ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuVal	96
Db	181	GTCACTTTGGGAGCCGGGCACTCGCTGTCCGCCAGGAGCCGTGCCACAGAGAGCTGTGTG	240
QY	97	AlaGluGlnAspGlnAspProSerGluLeuAsnProGlnThrGluGlnSerGlnAspPro	116
Db	241	GCAGAGAGAGACAGAGACCCGTCCGAACCTGAATCCCCAGACAGAAAGAAAGCCAGATCTT	300
QY	117	AlaProPheLeuAsnArgLeuValaIarGProArgArgSerAlaProIuysGlyArgIysThr	136
Db	301	GCGCCTTTCCTGAACCGGACTAGTTCGGCTCGGCGAATGGCAACCTTAAGGCCCGGAAACA	360
QY	137	ArgAlaIarGArgAlaIleAlaAlaIleIstArgIuValaIHisProAlaGProGluGlnAspGly	156
Db	361	CGGGCTCGAAGAGGCATCGGAGCCCATTTATGAAGTTCACGACCCCTGGAGAGAGCGA	420
QY	157	AlaGlnAlaGlyValaIAspGlyTThrValSerGlyTTPGluGlnAlaArgIleAsnSerSer	176
Db	421	CGCGAGGAGGTTGTGCAGAGGACACTGATGTGCTGGAGAGAAAGCCAGATATACAGCTCC	480
QY	177	SerProLeuIarGlyTArgAsnArgGlnIleGlyIuPheIleValTThrArgAlaGlyLeuTyr	196
Db	481	AGCCCTTCGCGCTCAACAGCCGACATCGGGAGATTATAGTCACCCCGGGCTGGGCTTAC	540
QY	197	TyrLeuTyrCysGlnValaIHisPheAspGluIuLysAlaValaIlyTyrLeuLysLeuAspLeu	216
Db	541	TACCTGTACTCTACAGGTGCATCTTTGATGATGAGGGAGAGGCTGTCTACTCTGACCTGAGCTTG	600

QY 217 LeuValAspGlyValIleuAlaIleuAlaArgCysLeuGluGluPheSerAlaThrAlaAlaSer 236

Db 601 CTrGGGGAGATGGTGTCTGTGGCCCTGGCGCTCCCTGGAGAGAAATTCTCAACCACCTGGCGGAGT 660

QY 237 SerLeuGlyProGluIleuAlaArgLeuCysGlnValAsrGlyLeuIleuAlaIleuAlaArgProGly 256

Db 661 TCCCTCGGGCCCCACCTCCCGCTTCGCCAGATGTCTGGGCTTTGGCCCTGGGGCCAGGG 720

QY 257 SerSerLeuAlaGlyIleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThr 276

Db 721 TCCCTCCCTGGCGGATCCGCCACCTCCCTCGGGGCCCATCTCAAGCGTGGCCCCCTCTCTCACCC 780

QY 277 TyrPheGlyLeuPheGluValHis 284

Db 781 TACTTCGGAGCTTCTCCAGGTTCA 804

RESULT 2	AF055872	1368 bp	mRNA	linear	PRI 04-MAY-1998
LOCUS	AF055872				
DEFINITION	Homo sapiens Apol3/DR3 11qand (Apol3) mRNA, complete cds.				
ACCESSION	AF055872				
VERSION	AF055872.1	GI:3108230			

CONTRIBUTOR	HOMO CAPTIVUS
REFERENCE	Euthariota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 1368)
AUTHORS	Marsters, S.A., Sheridan, J.P., Pitti, R.M., Brush, J., Goddard, A. and Ashkenazi, A.

TITLE	Identification of a ligand for the death-domain-containing receptor Apo3
JOURNAL	Curr. Biol. 8 (9), 525-528 (1998)

REFERENCE	2 (bases 1 to 1368)
AUTHORS	Marsters,S.A., Sheridan,J.P., Pittl,R.M., Brush,J., Goddard,A. and Ashkenazi,A.
TITLE	Direct Submission
JOURNAL	Submitted (25-MAR-1998) Molecular Oncology, Gentech, 1 DNA Way, South San Francisco, CA 94080, USA

FEATURES	Location/Qualifiers
source	1. .1368

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
1..1368
gene
/gene="APO3L"
58..807
CDS

```

```

/gene="APO3L"
/function="binds to the death-domain-containing receptor

```

Apo3/DR3<sup>a</sup>  
 /note="TNF homolog"  
 /codon\_start=1  
 /product="Apo3/DR3 ligand"  
 /protein\_id="AAC39724.1"  
 /db\_xref="GI:3108231"  
 /translation="MAARRQRRRRGRGRRGPTALVPLALGLGALACGLLAVV  
 GRRRSLSAQPAAEELVAEEDDOPSELNPTQTESQDPAFLNRLVNPRRSARKGR  
 ARRIAIAHYEHPRPDQAGQAOQAPLDTVSGMEARINSSPLTNMNOIDEEFTVTR  
 YLYYLCQVHFDEGRKAAVYLDLIDLVDGVLTALRCEEFSATAASSLGPOLRLCOVSGLLI  
 RPSSLSIRTLPPAHKLAAAFVLTGFLFOVH"  
 272 a 443 c 389 g 264 t

Alignment Scores:	
Pred. No.:	4,71e-63
Score:	1366.00
Matches:	268
Percent Similarity:	100.00%
Conservative:	0
Best Local Similarity:	100.00%
Mismatches:	0
Query Match:	94.60%
Indels:	0

DB:	9	Gaps:	0
US-09-245-198A-4 (1-284) x AF055872 (1-1368)			
QY 17 ArgSerLeuGlySerArgAspGlyValAlaValArgInalagInProProAlaBromet 36			
DB 1 CGATCCCTCGGGATCCCGGGATGGGGGGGGCGGTGAGGACAGGACACCCCGCCCGCCATTT 60			
QY 37 AlAlaAlaArgAspSerGlnArgArgGlyLyrArgArgGlyLupProGlyYThrAlaLeuLeu 56			
DB 61 GCCCCCGGTGGAGCCAGAGCGAGGGGGGGCCGGGGGGAGCCGGGACCGCTGTG 120			
QY 57 ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaVal 76			
DB 121 GTCCCGGCTGGGCGTGGGGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTG 180			
QY 77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnLupProAlaGlnGluLeuVal 96			
DB 181 GTCAAGTTGGGGAGCCGGGCACTCGCTGTCGCCGCCCGAGAGCTGCCACAGAGAGCTGGTG 240			
QY 97 AlagIngluAlaSpGlnAspProSerGluLeuAsnProGlnInlGluSerGlnAspPro 116			
DB 241 GCAGAGAGAGAGCCAGAGACCCCTCGGAATGATCCCGACAGAGAGAGAGAGAGAGAGCT 300			
QY 117 AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLysThr 136			
DB 301 GCGGCTTCCCTGGAACCGCACTAGTTCGGCTCGGCAAGATGCACTTAAAGCGCGGAAACA 360			
.QY 137 ArgAlaArgArgAlaAlaAlaAlaAlaHisTyrGluValHisProArgProGlyGlnAspGly 156			
DB 361 CGGGCTCGAAGAGGAGATGCAAGCCATTATGAACTCATCCAGACCTGGACAGGAGCGA 420			
QY 157 AlaGlnAlaGlyValAspGlyThrValSerGlyTyrGluGluAlaArgIleAsnSerSer 176			
DB 421 GCGCAGCAGAGGTGTGGACGGGACAGTGAATGGCTGGGAGGAAGCCAGATCAACAGCTCC 480			
QY 177 SerProLeuAsnGlyTyrAsnArgGlnIleGlyLupPheIleValThrArgAlaGlyLeuTyr 196			
DB 481 AGCCCTGCGCCCTCAACCCGACAGTCGGGGAGTTTAAAGTCAACCCGGGCGGGCTTCAC 540			
QY 197 TyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValAlaTyrLeuLysLeuAspLeu 216			
DB 541 TACCTGATCTGTCAAGGTGCTACATTGATAGAGGGAAGGCTGTCTACCTGAGCTGACTTG 600			
QY 217 LeuValAspGlyValLeuAlaLeuArgCysLeuGlnGluLupPheSerAlaThrAlaAser 236			
DB 601 CTGGTGATGTGTGTCTGGCCCTGGCTGGCGCTGGAGGAATCTCAGCCACTGCGCGAGT 660			
QY 237 SerLeuGlyProGlnInleuArgLeuGlyGlnValSerGlyLeuLeuAlaLeuArgProGly 256			
DB 661 TCCCTCGGGGCCCCAGCTCCGCTCTGCCCAAGGTGTGTGGGCTGTGGCCCTGGGGGCAAGG 720			
QY 257 SerSerLeuArgIleArgTyrThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThr 276			
DB 721 TCTCTCCGTCGGAGATCCGACACCTCCCTGGGGCCATCTCAAGAGCTGCCCCCTCTCTCAC 780			
QY 277 TyrPheGlyLeuPheGlnValHis 284			
DB 781 TACTTCGGAGCTCTTCACAGTTTCAC 804			

LOCUS	BC019047	1651 bp	mRNA	linear	PRI 11-DEC-2001
DEFINITION	Homo sapiens, similar to tumor necrosis factor (ligand) superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA, complete cds.				
ACCESSION	BC019047				
VERSION	BC019047.1	GI:17512138			
KEYWORDS	MGC.				
SOURCE	Homo sapiens.				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				

REFERENCE	1 (Phases 1 to 1651)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC Project URL: <a href="http://mgc.nci.nih.gov">http://mgc.nci.nih.gov</a>
COMMENT	Contact: MGC help desk Email: <a href="mailto:cgabs_email.nih.gov">cgabs_email.nih.gov</a> Tissue Procurement: Louis Staudt cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: <a href="http://www.nisc.nih.gov/nisc.mgcenhgri.nih.gov">http://www.nisc.nih.gov/nisc.mgcenhgri.nih.gov</a> Contact: <a href="mailto:shveychenko.y.,wetherby.r.d.,beckstrom.sternberg.s.m.,benjamin.b.,blakesley.r.w.,bouffard.g.,brinkley.c.,brooks.s.,dietrich.n.l.,guan.x.,gupta.j.,ho.s.-l.,karlins.e.,legaspi.r.,lim.m.,maduro.q.l.,masiello.c.,mastrian.s.d.,mccloskey.j.c.,mcdowell.j.,pearson.r.,snyder.b.,stantipop.s.,thomas.p.j.,tongsom.e.e.,touchman.j.w.,tsurgon.c.,vogt.j.l.,walker.m.a.,zhang.l.-h. and green.e.d.">shveychenko.y.,wetherby.r.d.,beckstrom.sternberg.s.m.,benjamin.b.,blakesley.r.w.,bouffard.g.,brinkley.c.,brooks.s.,dietrich.n.l.,guan.x.,gupta.j.,ho.s.-l.,karlins.e.,legaspi.r.,lim.m.,maduro.q.l.,masiello.c.,mastrian.s.d.,mccloskey.j.c.,mcdowell.j.,pearson.r.,snyder.b.,stantipop.s.,thomas.p.j.,tongsom.e.e.,touchman.j.w.,tsurgon.c.,vogt.j.l.,walker.m.a.,zhang.l.-h. and green.e.d.</a>
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Series: IRAL P.M.A.G.E. 30 Row: P Column: 5 This clone was selected for full length sequencing because it passed the following selection criteria: "matched mRNA g1: 4507596. Location/Qualifiers 1. 1651 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="MGC:20669 IMAGE:4766071" /tissue_type="Primary B-cells from Tonsils" /clone_id="NIH-MGC_48" /lab_host="DH10B-R" /note="Vector: pOTB7" 106..510 /product="Similar to tumor necrosis factor (ligand) superfamily, member 12" /protein_id="AAH19047.1" /db_xref="GI:17512139" /translation="MAARSRRRGRGEPGTALGLALACLGILAVVSLGSASISAOEPAQELVAAEDDPSSEINRGTETESQADAPPLNRYRRBRAPGRKTRARRAIAHYEYHVRPGDQADAGSGYTCCARP"
CDS	BASE COUNT 344 a 517 c 481 g 309 t ORIGIN
Alignment Scores:	
Pred. No.:	3 53e-61 Length: 1651
Score:	1331.50 Matches: 275
Percent Similarity:	77.25% Conservative: 0
Best local Similarity:	77.25% Mismatches: 1
Query Match:	92.21% Indels: 81
DB:	9 Gaps: 1
US-09-245-198A-4 (1-284) x BC019047 (1-1651)	
OY	9 SerAlaArgArgLeuProLeuProArgSerLeuGlySerArgAspGlyAlaValArg 28
Db	25 TCGCGCCGGCGGCTCCCTCCCGAGACCTCGGGTCCGGAGTGGGGGGCGGTAGG 84
OY	29 GlnAlaGlnProPheAlaProMetAlaAlaArgArgSerGlnArgArgGlyArgArg 48
Db	85 CAGGACAGAGCCCGCCCGCCATGAGCGCCGTCGAGACAGAGGCGAGGGGGCGCCGG 144
OY	49 GlyIleuProGlyThrAlaIleuLeuValProleuAlaIleuGlyIleuGlyLeuAlaIleu 68
Db	145 GGGGAGCGGGGACCGGCTCTGTGCTCCGCTCGCGGCGGCGCTGGGGCTGGCGC 204

```

QY      69  CysLeuGlyLeuLeuLeuAlaValSerLeuGlySerArgAlaSerLeuSerAlaGln 88
      205  TGGCTGGGCTCTGCTGGCGGTGCTGCTAGTTGGGGAGCGGGGATGCTGTCCGCCAG 264
QY      89  GluProAlaGlnGluLeuValAlaGluGlnAspGlnAspProSerGluLeuAsnPro 108
      265  GAGCTGCGCCAGAGAGAGCTGGTGCGAGAGAGACCGACCGCTCGGAATGAAATCCC 324
QY      109  GlnThrGluGlnSerGlnAspProAlaProPheLeuAsnArgLeuValArgProAlaArg 128
      325  CACACGAGAAAGAACCAAGATCTGCGCTCTTCTGTAACCGACTAGTTGGCTCGCAGA 384
QY      129  SerAlaProGlyGlyArgGlyThrArgAlaArgAlaIleAlaIleAspIleArgGluVal 148
      385  AGTGCACCTAAAGCGCGGAAACACAGCGCTCGAAGAGCGATGCGACCCATTATGAAGTT 444
QY      149  HisProArgProGlyGlnAspGlyAlaGlnAla----- 159
      445  CATCCAGACCTGGACAGAGACGAGCGAGCC-AGATGGAGTTACACAACATTGTCTGAG 503
QY      159  ----- 159
      504  GCCATGATGACTAGTGGTGGAGCCAAAGATTGAAACCAAGTAAATGTGCTGTAC 563
QY      159  ----- 159
      564  TGGACATGCTGTGATGAAGAGCGAGGTGCTGTGTCAGGGGTGAGGGGTCCATGCAGGG 623
QY      159  ----- 159
      624  GCCACATCCAAAAGGGGAGAGGAAATTCCAGAAAGAGAGACATCTCCACCC 683
QY      160  -----GlyValAspGlyThrValSerGlyTrp 168
      684  ATTACAGAGGCTCAAGAGAGAACAGAGACCAAGCTGTGGAGGGACATGATGGCTGG 743
QY      169  GluGluAlaArgIleAsnSerSerProLeuArgTyraAsnArgGlnIleGlyGluPhe 188
      744  GAGGAAGCCAGAAATCAACAGCTCCAGCCCTGTGCGCTACAAACGCCAGATCGGGAGTTT 803
QY      189  IleValThrArgAlaGlyLeuTyrrTyrrLeuTyrrCysGlnValHisPheAspGluGly 208
      804  ATAGTCACCGGGCTGGCTGTACTACTGTCTGTCAGTGTCTGCTGATGAGGGAG 863
QY      209  AlaValTyrrLeuLeuAspLeuValAspGlyValIleLeuAlaLeuArgGlySerGlu 228
      864  GCTGTCTACTGTAAGCTGGACTGCTGTGGATGTGTGCTGGCTGGCTGGCTGGCTGG 923
QY      229  GluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgLeuGlyGlnValSer 248
      924  GAATTCCTAGCCACTGCGCGGCTTCCTCGGGGCCCAAGCTCGGCTCTGCCAGGTGCT 983
QY      249  GlyLeuLeuAlaLeuArgProGlySerSerSerLeuArgIleArgThrLeuProTrpAlaHis 268
      984  GGGCTGTGGCTGGCGGCGGAGGCTCTCCCTGCGGATCCACACCTCCCTCGGGCCCAT 1043
QY      269  LeuTyraAlaProPheLeuThrTyrrPheGlyLeuPheGlnValHis 284
      1044  CTCAGAGCTGCCCTTCCTCACCCTACTTGGGACTCTTCCAGTTCCAC 1091

```

RESULT 4  
 LOCUS AF030099 1306 bp mRNA linear PRI 20-DEC-1997  
 DEFINITION Homo sapiens TWEAK mRNA, complete cds.  
 ACCESSION AF030099  
 VERSION AF030099.1 GI:2707218  
 KEYWORDS  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1306)

```

AUTHORS Chicheportiche, Y., Bourdon, P.R., Xu, H., Hsu, Y.M., Scott, H.,
        Hession, C., Garcia, I. and Browning, J.U.
TITLE TWEAK, a new secreted ligand in the tumor necrosis factor family
        that weakly induces apoptosis
JOURNAL J. Biol. Chem. 272 (51), 32401-32410 (1997)
MEDLINE 98070415
PUBMED 9405449
REFERENCE 2 (bases 1 to 1306)
AUTHORS Bourdon, P., Hession, C., Tizard, R. and Browning, J.
TITLE Direct Submission
JOURNAL Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,
        Cambridge, MA 02142, USA
FEATURES
source Location/Qualifiers
    1..1306
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="17"
        /map="17p13"
        /tissue_type="consil"
        18..767
        /note="Ligand in the TNF family; secreted protein; start
        codon not verified experimentally"
        /product="TWEAK"
        /protein_id="AAC51923.1"
        /db_xref="GI:2707219"
        /translation="MAARRSORRRGRGEPGATLVLPALGIALAGLAVSL
        GSRASLSAEPDAEELVAEEDDPELNPQTEESDDPAPFLNVRPRSPKGRKT
        ARRAIAAHYEVHPRGDQAGAVGTSGVSGEARSINSPLRVNROIGETIVRAGI
        YLVGVHFEDEGKAVYLRKLDLVGVSLALRCLERSARAAASLSLPQLCOVSGILAL
        RPSGLRTRTPMHLKAPFLTYGTFQVH"
BASE COUNT 247 a 434 c 368 g 257 t
ORIGIN
Alignment Scores:
Pred. No.: 1,53e-59 Length: 1306
Score: 1298.00 Matches: 254
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.89% Indels: 0
DB: 9 Gaps: 0
US-09-245-198A-4 (1-284) x AF030099 (1-1306)
QY      31  GlnProProAlaProMetAlaAlaArgArgSerGlnArgArgGlyArgGlyGlu 50
      3  CAGCCCCCGCCCCCATGGCGCCGCTGGAGCCAGAGCGGAGGGCGCGGGGGAG 62
QY      51  ProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeu 70
      63  CCGGACCGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 122
QY      71  GlyLeuLeuLeuAlaValAlaSerLeuGlySerArgAlaSerLeuSerAlaGlnGluPro 90
      123  GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 182
QY      91  AlaGlnGluGluLeuValAlaGluGlnAspGlnAspProSerGluLeuAsnProGlnThr 110
      183  GCCCAGGAGGAGCTGGTGCGAGAGAGACCAAGACCCGCTGGAATCCCAATCCCAAGCA 242
QY      111  GluGlnSerGlnAspProAlaProPheLeuAsnArgLeuValArgProAlaArgSerAla 130
      243  GAAGAAAGCCAGAGATCTGCGCTTCTTGAAACGACTAGTTGGGCTCGCGCAAGTGA 302
QY      131  ProGlyGlyArgGlyThrArgAlaArgAlaIleAlaIleAspIleArgGluValHisPro 150
      303  CCTAAAGCGCGGAAACACAGGGCTCGAAGAGCGATCCAGCCCTTATGAAGTTCAATCCA 362
QY      151  ArgProGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTrpGluGlu 170
      363  CGACCTGGACGAGAGAGAGCGCAGGAGAGTGTGAGACGGACAGAGTGTGCTGGAGGAA 422
QY      171  AlaArgIleAsnSerSerProLeuArgTyrrAsnArgGlnIleGlyGluPheIleVal 190

```



```

|||||
Db 423 GCCAGATACACGCTCCGACCCCTCGCTACACCGCCAGATCGGGAGTTTATAGTC 482
Oy 191 ThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGluIlylsAlaVal 210
Db 483 ACCCGGGCTGCTACTACTGTAAGTGTACAGTGTGATGATGAGGGAAGGCTGTC 542
Oy 211 TyrLeuIlyLeuAspLeuLeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPhe 230
Db 543 TACCTGAAGCTGACTGCTGTGGATGGTGTGCTGGCCCTGCGCTCGCTGGAGGAATTC 602
Oy 231 SerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeu 250
Db 603 TCAGGCACATGGCGGACGATTCCTCGGGGCCACAGCTCCGCTGCTGCGAGGTCTGGGCTG 662
Oy 251 LeuAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProThrAlaHisLeuLys 270
Db 663 TTGGCCCTGCGGCGCAGGGTCTCCCTGCGGATCCGACCCCTCGGCGCCATCTCAAG 722
Oy 271 AlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 284
Db 723 GCTGCCCTTCTCTCACTTCTGAGACTTCTTCAGGTTTCAC 764

RESULT 5
AR140407 1236 bp DNA linear PAT 16-JUN-2001
LOCUS AR140407 Sequence 1 from patent US 6207642.
ACCESSION AR140407.1 GI:14482903
VERSION AR140407.1 GI:14482903
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1236)
AUTHORS Wiley,S.R.
TITLE Member of the TNF family useful for treatment and diagnosis of
JOURNAL Patent: US 6207642-A 1 27-MAR-2001;
FEATURES
source 1.1236
location/Qualifiers
BASE COUNT 225 a 416 c 358 g 237 t
ORIGIN

Alignment Scores:
Pred. No.: 7.51e-58 Length: 1236
Score: 1265.00 Matches: 248
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.60% Mismatches: 0
Query Match: 87.60% Indels: 0
Gaps: 0
DB: 6

US-09-245-198a-4 (1-284) x AR140407 (1-1236)
Oy 36 MetAlaAlaArgSerGlnArgArgArgGlyArgGlyGluProGlyThrAlaLeu 55
Db 1 ATGGCGCGCCGCTCGAGGAGCAAGCGAGGGGCGCGCGGAGCGCGGCGCGCTG 60
Oy 56 LeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAla 75
Db 61 CTGGTCCCGCTCGCGCTGGGCGCTGGGCGCTGGCTGCGCTGCGCTGCTGCTGCG 120
Oy 76 ValValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluLeu 95
Db 121 GTGGTCTAGTTTGGGAGCGCGGCGATCGTCTCGCGCCAGAGAGCTCCAGAGGAGCTG 180
Oy 96 ValAlaGluGluAspGlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAsp 115
Db 181 GTGGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 240
Oy 116 ProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLys 135
Db 241 CCGGCGCTTCTCTGAGACGAGACTAGTTGCGGCTCGAAGAAAGTGCACCTAAAGCGCGA 300

```

```

Oy 136 ThrArgAlaArgArgAlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAsp 155
Db 301 ACACGGGCTCGAAGACGATCGCAGCCCATTTATGAAGTTTCATCCAGACCTGGACAGGAC 360
Oy 156 GlyAlaGluAlaGlyValAspGlyThrValSerGlyTyrGluGluValArgIleAsnSer 175
Db 361 GAGAGCCAGGAGGTGTGACGAGGACAGTGAAGTGGCTGGGAGGAGACCAATCAACAGC 420
Oy 176 SerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeu 195
Db 421 TCAGCGCTCTGCGCTACACAGCGCAGATCGGGAGGATTTATAGTACACCGGCGCTGGGCTC 480
Oy 196 TyrTyrLeuTyrCysGlnValHisPheAspGluGlyValAlaValTyrLeuLysLeuAsp 215
Db 481 TACTACCTTACTGTAAGTGTGACCTTGTGAGAGGAGAGGCTGCTTACTGAAGCTGGAC 540
Oy 216 LeuLeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAla 235
Db 541 TTGCTGTGATGATGTGTGCTGGCGCTGCGCTGCTGAGAGAAATTCACAGCACTGGCGCG 600
Oy 236 SerSerLeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgPro 255
Db 601 AGTTCCCTGGGCGCCAGCTCCGCTGCTGCGAGGTGTGTGGGCTGTGGCGCTGCGGCA 660
Oy 256 GlySerSerLeuArgIleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeu 275
Db 661 GGGTCTCTCCCTGGGATCGGACCCCTCGGCGCCATCTCAAGCTGCGCCCTTCTCTC 720
Oy 276 ThrTyrPheGlyLeuPheGlnValHis 284
Db 721 ACTTACTTGGAGCTTCTCCAGGTTTCAC 747

RESULT 6
AX180714 898 bp DNA linear PAT 06-AUG-2001
LOCUS AX180714 Sequence 1 from Patent WO0145730.
ACCESSION AX180714
VERSION AX180714.1 GI:15132570
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 898)
AUTHORS Wiley,S.R.
TITLE Tweak receptor
JOURNAL Patent: WO 0145730-A 1 28-JUN-2001;
FEATURES
source 1.898
location/Qualifiers
CDS
/note="human TWEAK fusion protein construct"
/db_xref="taxon:32630"
52..873
/note="human TWEAK fusion protein construct"
/translation="1"
/protein_id="CAC50004.1"
/db_xref="GI:15132571"
/translation="MATGSRITLLAFGLICLPMTLOEGSATSDDMKQIEDKIEILS
KYVHIENEIARIKLIGERTRSLSRASLSAOEPAQELVAEEDQDSLEINPTEES
ODPAFLENLVRPERSAPRGKTRARBAIAAYEVHPRGODGAGADGTVSGMEEA
RINSSPLRNVROIGEFVTRAGIGLYIYICQVAFDEGKANYITLIDLINCVALRCL
FSATASSTIGPOLRICOVSGIALRPGSSLRIRITLPMALIKAPLITYFGLRQVH"
BASE COUNT 187 a 266 c 267 g 178 t
ORIGIN

Alignment Scores:
Pred. No.: 1.19e-47 Length: 898
Score: 1066.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.82% Indels: 0
Gaps: 0
DB: 6

```

US-09-245-198a-4 (1-284) x AX180714 (1-898)

QY 78 SerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluValAla 97  
 |||||||  
 DB 250 AGTTTGGGACCCGGGCTCTGCTGCCCGACAGAGCTGCCAGAGAGAGCTGGTGGCA 309

QY 98 GluGluAspGlnAspProSerGluLeuAsnProGlnTrpGluGluSerGlnAspProAla 117  
 |||||||  
 DB 310 GAGAGGACGACAGACCCCTCGGAAGTGAATCCCGACAGAGAAAGCCAGAGATCTCGC 369

QY 118 ProHeLeuAsnArgLeuValAlaGProArgArgSerAlaProGlyGlyArgGlyThrArg 137  
 |||||||  
 DB 370 CTTTCGGAACGACGACTGATGCTGCGCTCGAGAGTGCACCTAAAGCCGGGAAACACAGG 429

QY 138 AlaArgArgAlaLeuAlaAlaHisTrpGluValHisProArgProGlyGlnAspGlyAla 157  
 |||||||  
 DB 430 GCTGAGAGAGCGGATCGCGCCCATTTATGAATTCATCCACAGCTCGGACAGAGCGAGCG 489

QY 158 GlnAlaGlyValAlaAspGlyThrValSerGlyTrpGluGluAlaArgGlnAsnSerSer 177  
 |||||||  
 DB 490 CAGGACAGGTGTGGACGAGGACAGTGAAGTGGGAGGAGAAAGCAATCAACAGCTCCAGC 549

QY 178 ProLeuArgTrpAsnArgGlnIleGlyGluPheIleValIleThrArgAlaGlyLeuTyrTyr 197  
 |||||||  
 DB 550 CCTGCGCCTACAAACCGCCAGATCGGGAGTTTATATGACCCCGGCTGCTACTAC 609

QY 198 LeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu 217  
 |||||||  
 DB 610 CTGACTGTCAAGTGCATTTGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 669

QY 218 ValAspGlyValLeuAlaLeuValGlyLeuGluGluPheSerAlaThrAlaAlaSerSer 237  
 |||||||  
 DB 670 GTGATGTGTGCTGGGCTGCTGCTGCTGAGGAAATCTCAGCCACTGCGGCCAGTTCC 729

QY 238 LeuGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySer 257  
 |||||||  
 DB 730 CTCGGGCCCCAGCTCCGCTCTGCCAGGTGTCTGGGCTGTGGCCCTCGGCGAGGATCC 789

QY 258 SerLeuArgGlnArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyr 277  
 |||||||  
 DB 790 TCCCTGGGATCCGACCCCTCCCGGGGCCATCATCAAGGCTGCCCTCTCTACCTAC 849

QY 278 PheGlyLeuPheGlnValHis 284  
 |||||||  
 DB 850 TTCGACTCTTCAGGTTCCAC 870

RESULT 7  
 AF030100 1168 bp mRNA linear ROD 20-DEC-1997  
 LOCUS Mus musculus TWEAK mRNA, partial cds.  
 DEFINITION AF030100  
 VERSION AF030100.1 GI:2707220  
 KEYWORDS  
 SOURCE Mus musculus.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1168)  
 Chicheportiche Y., Bourdon, P.R., Xu, H., Hsu, Y.M., Scott, H.,  
 Hession, C., Garcia, I. and Browning, J.L.  
 TWEAK, a new secreted ligand in the tumor necrosis factor family  
 that weakly induces apoptosis  
 J. Biol. Chem. 272 (51), 32401-32410 (1997)  
 9405449  
 PUBMED 2 (bases 1 to 1168)  
 Chicheportiche Y., Bixler, S., Tizard, R. and Browning, J.  
 Direct Substitution  
 Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,  
 Cambridge, MA 02142, USA  
 FEATURES  
 source 1..1168

CDS

/organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /cell\_type="peritoneal macrophages"  
 <1..679  
 /note="Ligand in the TNF family; secreted protein"

BASE COUNT 242 a 360 c 298 g 268 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3,79e-45 Length: 1168  
 Score: 1020.00 Matches: 199  
 Percent Similarity: 92.86% Conservative: 9  
 Best Local Similarity: 88.84% Mismatches: 16  
 Query Match: 70.64% Indels: 0  
 DB: 10 Gaps: 0

US-09-245-198a-4 (1-284) x AF030100 (1-1168)

QY 61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaValSerLeuGly 80  
 |||||||  
 DB 5 CTGAGCCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 64

QY 81 SerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluValAlaGluGluAsp 100  
 |||||||  
 DB 65 AGCTGGGCAAGCAGCTGTGCCAGAGAGCTCTCAGAGAGAGCTGACAGAGAGAGAGAGAGAG 124

QY 101 GlnAspProSerGluLeuAsnProGlnTrpGluGluSerGlnAspProAlaProPheLeu 120  
 |||||||  
 DB 125 CGGAGCCCTGGAATCAATCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 184

QY 121 AsnArgLeuValArgProArgArgSerAlaProGlyGlyAlaGlyThrArgAlaArg 140  
 |||||||  
 DB 185 GACACACTAGTCCGCGCTCGAAGAGTCTCTTAAGGCCGAGAGAGAGAGAGAGAGAGAGAGAG 244

QY 141 AlaIleAlaAlaHisTrpGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 160  
 |||||||  
 DB 245 GCTATTGACGCCCATTTATGAGGTTCATCTCGGCGAGAGAGAGAGAGAGAGAGAGAGAGAG 304

QY 161 ValAspGlyThrValSerGlyTrpGluGluAlaArgGlnAsnSerSerProLeuArg 180  
 |||||||  
 DB 305 GTGATGGGACAGTAGTGGCTGGGAGAGACCAAAATCAACAGCTCCAGCCCTCTGGCGC 364

QY 181 TyrAsnArgGlnIleGlyGluPheIleValIleThrArgAlaGlyLeuTyrTyrLeuTyrCys 200  
 |||||||  
 DB 365 TACGACCCGCAAGTTGGGGAATTTACAGTCACTACAGGCTGGGCTCTACTACTGACTGCT 424

QY 201 GlnValHisPheAspGlyGlyLysAlaValTyrLeuLysLeuAspLeuValAspGly 220  
 |||||||  
 DB 425 CAGGACACTTGTGATGAGGAGAAAGGCTGTCACTGAAGCTGACTGCTGGTGAACGCT 484

QY 221 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerLeuGlyPro 240  
 |||||||  
 DB 485 GTGCGGCTGGGCTGGGCTGGGAGAAATTCACGACAGACAGCAAGAGTCTCTGGGCGC 544

QY 241 GlnLeuArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArg 260  
 |||||||  
 DB 545 CAGCTCCGTTGTGCGAGAGTGTGGGCTGTGGCGCGCTGGGCGAGAGGCTCTCCCTTCGCG 604

QY 261 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 280  
 |||||||  
 DB 605 ATCCGACACCTCCCTGGGCTGATGAGCTGAGGCTGGCCCTTCTTAACCTACTTGGAGCTC 664

QY 281 PheGlnValHis 284  
 |||||||

Db 665 TTFCAGTTCAC 676

RESULT 8  
AC016876/c  
LOCUS  
DEFINITION  
Homo sapiens clone RP11-186B7, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 10  
UNORDERED PIECES.

AC016876  
AC016876  
AC016876.5 GI:21313830  
HTG; HTGS\_PHASE1; HTGS\_FULLTROP; HTGS\_ACTIVEFIN.  
human.

ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

1 (bases 1 to 177703)  
Bliren,B., Nusbbaum,C. and Lander,E.  
Homo sapiens chromosome, clone RP11-186B7  
Unpublished  
2 (bases 1 to 177703)  
Bliren,B., Linton,L., Nusbbaum,C., Lander,E., Allen,N., Anderson,M.,  
Baldwin,J., Barna,N., Beckery,R., Boguslavsky,L., Boukhalter,B.,  
Brown,A., Castle,A., Collange,R., Collins,S., Collymore,A.,  
Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,  
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,  
Galagan,J., Gardyna,S., Grant,G., Hages,B., Hatford,A., Horton,L.,  
Howard,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Lenocksky,J., Lieu,C., Locke,K., MacDonald,P., Marquis,N.,  
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,  
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
Wymann,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles street, Cambridge, MA 02141, USA  
3 (bases 1 to 177703)  
Bliren,B., Nusbbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,  
Camarta,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,  
Cooke,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hages,B.,  
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,  
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,  
Liu,G., Maclean,C., MacDonald,P., Major,J., Matthews,C.,  
McCarthy,M., Meldrim,J., Menes,L., Mihova,T., Mieng,V.,  
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,  
Plunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,  
Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P.,  
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,  
Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wymann,D., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 2, 2002 this sequence version replaced gi:15421989.  
All repeats were identified using RepeatMasker:  
Smt, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WtBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L3849  
Center clone name: 186\_B\_7

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 10 contigs. The true order of the pieces

	* is not known and their order in this sequence record is
	* arbitrary. Gaps between the contigs are represented as
	* runs of N, but the exact sizes of the gaps are unknown.
	* This record will be updated with the finished sequence
	* as soon as it is available and the accession number will
	* be preserved.
	1. 48645: contig of 48645 bp in length
	* 48646 48745: gap of 100 bp
	* 48746 58277: contig of 9532 bp in length
	* 58278 58377: gap of 100 bp
	* 58378 65804: contig of 7427 bp in length
	* 65805 65904: gap of 100 bp
	* 65905 79793: contig of 13879 bp in length
	* 79794 79893: gap of 100 bp
	* 79894 99493: contig of 19600 bp in length
	* 99494 99593: gap of 100 bp
	* 99594 111049: contig of 11456 bp in length
	* 111050 111149: gap of 100 bp
	* 111150 125020: contig of 13871 bp in length
	* 125021 125120: gap of 100 bp
	* 125121 145109: contig of 19989 bp in length
	* 145110 145209: gap of 100 bp
	* 145210 169458: contig of 24249 bp in length
	* 169459 169558: gap of 100 bp
	* 169559 177703: contig of 8145 bp in length.
FEATURES	Location/Qualifiers
source	1..177703
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/clone_id="RP11-186B7"
BASE COUNT	42862 a 45967 c 46568 g 41254 t 1052 others
ORIGIN	
Alignment Scores:	
Pred. No.:	1.39e-20 Length: 177703
Score:	589.50 Matches: 125
Percent Similarity:	65.62% Conservative: 1
Best Local Similarity:	65.10% Mismatches: 0
Query Match:	40.82% Indels: 66
	Gaps: 1
DB:	
US-09-245-198A-4 (1-284) x AC018676 (1-177703)	
OY 159 AlaglyValaspLeuThyValSerGlyTyrPglLualArgIleasnSerSerPro 178	:::
Db 37636 TCAGGTGGTGACGCCGACACTGAATGGCGGAGCAACCACAATCAGCTCACGCCCT 37577	
OY 179 LeuarGtTyraArGlnIleelGylupheIleValThrArGalaglLeuTyrTyrlau 198	
Db 37576 CTCGGCTACAAACCCGCAGATCGGGAGTTATATAGCACCCGGGCTCTACTACCTG 37511	
OY 199 TyrCys----- 200	
Db 37516 TACTGTCAAGTAAGCCCCCATCTGGCTGCATGGGTAAACGAGTAAGAAGATGGCGAGGCT 37457	
OY 200 ----- 200	
Db 37456 TTGCCAGAGAGACGTGGGGGACAAAGCTACACAGCGCTGGGAGAGTGAGTGGGGTTTTGGGTGGG 37397	
OY 200 ----- 200	
Db 37396 ATGGAGATGCCCTGCTGCTGAGGAATTGGAATAATGAGCGAGGGACAGGAGAGCCCTGG 37333	
OY 201 ----- GlnValHispheAspIuclLylysAlvalTyrlau 213	
Db 37336 ACTCGGCCTGTGTTCGCCACCCACAGGTGCACCTTATATAGGGGAGAGCTGTCTAACCTGAA 37277	
OY 213 sleaAspLeuValaspLlValIleuaLeuArGcysLeuGlupheserAlarh 233	
Db 37276 GCTGGACTTCGTCTGTGTGGATGTGTGTCTGCGCTGGCTGGAGCAATTCACGCCAC 37211	
OY 233 ralaAlaserSerleuGlyProGluLeuArGcysGlnValSerGlyleuAlate 253	





clone\_end:SP6  
vector\_side:right"  
BASE COUNT 41026 a 46997 c 46681 g 41924 t 927 others  
ORIGIN

Alignment Scores:  
Pred. No.: 6.8e-19 Length: 177555  
Score: 557.00 Matches: 130  
Percent Similarity: 52.82% Conservative: 20  
Best Local Similarity: 45.77% Mismatches: 45  
Query Match: 38.57% Indels: 89  
DB: 2 Gaps: 4

US-09-245-198a-4 (1-284) x AC130192 (1-177555)

67 LeuAlaCysLeuGlyLeuLeuAlaValValSerLeuGlySerArgAlaSerLeuSer 86  
DB 114760 TTGCTCTGTAATCTTTCTGTCATGTAATGTTGGATTCTTTCACCTCATTCTATCACA 114701  
87 AlaGlnGluPro-----AlaGlnGluGluLeuValAlaGluGluAspGlnAspPro 103  
DB 114700 CTTTGGGCTCCAGAGCCGATGCTCCCTTCCTTCAGCCATCTGTATCATCCG 114641  
104 SerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeu 123  
DB 114640 GGTACGGTGGCCCAAGGTTGATCAATTCAGTGGCCCGGGGATTATTA----- 114590  
124 ValArgProArgArgSerAlaProLysGlyArgLysThrArgAlaArgAlaIleAla 143  
DB 114589 -----GAAACCTCTGTAAGGACTTCC----- 114566  
144 AlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyAlaAspGly 163  
DB 114565 -----CCTGTCTCTGACATCCACTCAGCTGTAAGCGGG 114530  
164 ThrValSerLysTyrPglGluAlaArgIleAsnSerSerProLeuArgTyrAsnArg 183  
DB 114529 ACGGTGAGTGGTGGGAGGAGGCAAAATCAACAGCTCCACACCTGCTATGACCGC 114470  
184 GlnIleGlyLupPheIleValIleThrArgAlaGlyLeuTyrTyrLeuTyrCys----- 200  
DB 114469 CAGAGCGGGCAATTTTGGTGCACCGGGGAGGCTGTACTGCTCTCAGTAAAGC 114410  
200 ----- 200  
DB 114409 CCCACCTGGCTCCGCGGTGAGCAGAGACAGGAGAGAGGTTCTGGGGGGTGGGCT 114350  
200 ----- 200  
DB 114349 GGGCCCAAGTTAAAGTGGGAGAGGAGGAGTTGTGCTGGAAGATGCCGGGCTCTT 114290  
201 -----G1 201  
DB 114289 GGGAAACCTGAGACAGGAGGCAAGGAGCAGAGAGGCTGAGCTCCGCCGCCGCCCA 114230  
201 nValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGlyVal 221  
DB 114229 GGTGCACTTTGATGAGGAGGAGGCGCTCTACCTGAAGCTGACTGCTGTGGAACGACAC 114170  
221 LeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGluProG1 241  
DB 114169 GGTGGCCCTGCTGCTGAGAGAGTCTGCGACGCGGCGAGTCCCTCGGGGCCCA 114110  
241 nLeuArgLeuGlyGlnValSerLysLeuLeuAlaLeuArgProGlySerSerLeuArgI1 261  
DB 114109 GCTCCGCTCTGCGCAGGTGTCTGGGCTATTTGCCCTGCGGCGGCTCTCTCGCGGAT 114050  
261 eArgThrLeuProTyrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuP 281  
DB 114049 CCGCACCTCTCCCTGGGCGCACCTCAAGGCTGCGCCCTCTCACCCTTGTGAGACTCT 113990  
281 eGlnValHis 284  
|||||

DB 113989 CCAAGTTCAC 113980

RESULT 11  
AC126925/c  
LOCUS  
DEFINITION  
Canis familiaris clone RP81-332E11, WORKING DRAFT SEQUENCE, 15  
unordered pieces.  
AC126925  
AC126925.1 GI:21724102  
VERSION  
KEYWORDS  
HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE  
ORGANISM  
Canis familiaris  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
1 (bases 1 to 161428)

# REFERENCE

Akhter N., Antonellis A., Ayala K., Beckstrom-Sternberg S.M.,  
Benjamin B., Blakesley R.W., Bouffard G.G., Breen K., Brinkley C.,  
Brooks S., Dietrich N.L., Granite S., Guan X., Gupta D.,  
Haghighi P., Hansen N., Ho S.-L., Idol J.R., Karlins E., Latic P.,  
Lee-Hin S.-O., Legaspi R., Maduro O.L., Maduro V.B.,  
Margulies E.H., Mastello C., Maskeri B., Mastrian S.D.,  
McCluskey J.C., McDowell J., Pegurigan C., Pearson R.,  
Portnoy M.E., Prasad A., Schueler M.G., Stantiridop S., Thomas J.W.,  
Thomas P.J., Touchman J.W., Tsugeon C., Vogt J.L., Walker M.A.,  
Wetherby K.D., Wiggins L., Young A., Zhang L.-H. and Green E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 161428)  
Direct Submission  
Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717  
Groveomont Circle, Gaithersburg, MD 20877, USA  
----- Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: http://www.nisc.nih.gov  
Contact: nisc-zoo@ngt.nih.gov  
----- Project Information  
Center project name: 332E11  
Center clone name: 332E11  
----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 148712 bases at least Q40  
Consensus quality: 151471 bases at least Q30  
Consensus quality: 153125 bases at least Q20  
Insert size: 15200; agarose-fp  
Insert size: 160028; sum-of-coverage  
Quality coverage: 5.36x in Q20 bases; sum-of-coverage  
Quality coverage: 5.10x in Q20 bases; sum-of-coverage

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 15 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 4734: contig of 4734 bp in length  
\* 4735 4834: gap of unknown length  
\* 4835 7325: contig of 2491 bp in length  
\* 7326 7425: gap of unknown length  
\* 7426 12261: contig of 4836 bp in length  
\* 12262 12361: gap of unknown length  
\* 12362 19031: contig of 6670 bp in length  
\* 19032 19131: gap of unknown length  
\* 19132 27849: contig of 8718 bp in length  
\* 27850 27950: gap of unknown length  
\* 27950 31801: contig of 3852 bp in length  
\* 31802 31901: gap of unknown length  
\* 31902 40781: contig of 8880 bp in length



	Db	43125	CTGGCAAGGATGGGTGGGAGAGGAGACTCGCCACATCATGAAGAATGCCGCTT	43066
OY	201	-----	glnValHisPheas	205
Db	43065	TGCATGAGGAGCAGAGAGTCGTGATTGCTTGCCCTCTCTCTGTTCCAGAGTCACATTGGA	43006	
OY	205	PGLUGLYYSAlaValTYrLeuLYsLeuAspLeuValaspGlyValAlaLeuAr	225	
Db	43005	TGAGGAAAGGCTGCTCTACCTGAAAGCTGACTGCTGTGGTAAGGCTGTGCCCCCTGGC	42946	
OY	225	gcysLeuGIuGlunPheSerIarThrAlaLaserSerLeuGIyProGIuLeuArgLeucY	245	
Db	42945	CTGCCTGGAAAAATTCCTCAGCACAGCAGCAAGCTCTCTCGGGCCCCCAGACTCCGTTGGG	42886	
OY	245	sglValSerGIyLeuLeuAlaLeuArgProGIySerSerLeuAagIIeArgThleuPr	265	
Db	42885	CCAGGTGCTGGGCTGTGGCCGCTGGCCAGAGGCTTCCTTCGAGATCCGACCTCCC	42826	
OY	265	OTPAHIsHisLeuYSAlaAlaIaProPheLeuThTyrrPheGIyLeuPheGlnValHis	284	
Db	42825	CTGGGCTCATCTTAAAGCTGGCCCTCTCTCAACCTTAGACCTTTCAGATTAC	42768	
RESULT 13				
AL603707/c				
LOCUS		234182 bp	DNA	linear
DEFINITION		Mouse DNA sequence from clone RP23-422L16 on chromosome 11,		
ACCESSION		AL603707		
VERSION		AL603707.5	GI:17017790	
KEYWORDS		HFG.		
SOURCE		house mouse.		
ORGANISM		Mus musculus		
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE		1 (bases 1 to 234182).		
AUTHORS		Pearce,A.		
TITLE		Direct Submission		
JOURNAL		Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Nov 20, 2001 this sequence version replaced gi:16605765. During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations clones. The variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sr., SWISSPROT; Tr., TrEMBL; Wp., WormPeP; Information on the WormPeP database can be found at <a href="http://www.sanger.ac.uk/Projects/C_elegans/wormpep">http://www.sanger.ac.uk/Projects/C_elegans/wormpep</a> RP23-422L16 is from the RPC1-23 Mouse PAC library constructed by the group of Pleter de Jong. For further details see <a href="http://www.chori.org/bacpac/home.htm">http://www.chori.org/bacpac/home.htm</a> VECTOR: pBac3.6 This sequence is the entire insert of clone RP23-422L16.		
FEATURES		location/Qualifiers		
SOURCE		1..234182		
		/organism="Mus musculus"		
		/db_xref="taxon:10090"		
		/chromosome="11"		
		/clone="RP23-422L16"		
		/clone_id="RPC1-23"		
		complement(84050..84131)		





```

BASE COUNT      39178 a 39259 t 1107 others
ORIGIN
misc_feature    /note="assembly_fragment"
                8536. 15799
/note="assembly_fragment"
misc_feature    /note="assembly_fragment"
                15900. 25224
/note="assembly_fragment"
misc_feature    /note="assembly_fragment"
                25335. 32504
/note="assembly_fragment"
misc_feature    /note="assembly_fragment"
                32605. 40970
/note="assembly_fragment"
misc_feature    /note="assembly_fragment"
                41071. 56590
/note="assembly_fragment"
misc_feature    /note="assembly_fragment"
                56691. 73769
/note="assembly_fragment"
misc_feature    /note="assembly_fragment"
                73870. 90859
/note="assembly_fragment"
misc_feature    /note="assembly_fragment"
                90960. 111428
/note="assembly_fragment"
misc_feature    /note="assembly_fragment"
                111529. 153553
/note="assembly_fragment
clone_end:SP6
vector_size:right"

```

Alignment Scores:	
Pred. No.:	9.77e-16
Score:	495.00
Percent Similarity:	58.59%
Best Local Similarity:	55.56%
Query Match:	34.28%
DB:	2
	Gaps: 1

QY	152	ProGluGlnAspGlyValAlaGlnAlaGlyValAlaSpGlyThrValSerGlyTrpGluGlnVal	171
Db	26848	CCTGCTCTGTAAACACCCACATTCACAGTGTGACGGGACGGTGAAGTGGCTGGAGGAGGCC	26907
QY	172	ArgIleasnSerSerSerProLeuArgTyrAsnArgIlnIleGlyGluPheIleValThr	191
Db	26908	AAATTCACACACTCCACACCCCTCGCGTATGTACCTCCAGACCGGCAATTTCAGTCCACC	26967
QY	192	ArgAlaGlyLeuArgTyrTyrLeuTyrCysGlnVal	202
Db	26968	CGGGCTGGGCTGTACTACTCTACTACTGTACAGT-AAAGCCACACTGGCTCCACGGGTAAAG	27036
QY	202	-----	202
Db	27027	CCGGACACTGAAGAGAAAGCTGGCGCTTCGGGGGTCGGGGCCAGATTAAAGTGGGAGAGG	27086
QY	202	-----	202
Db	27087	GAGCGTGGGGTTTGGGCTGAGAGAGACCTTGGGCTCTAAGGACACCTGAGATGAAAGCC	27146
QY	203	-----HisPheAspGlu	206
Db	27147	AGGGCCACAGAGGCTTGACTCGGCCCTCCCTCCCTGCCCCCAGGTGCACATTTGATGAG	27206
QY	207	GlyIysAlaValIlyrLeuLysLeuAspLeuLeuValAspGlyValIleuAlaLeuArgCys	226
Db	27207	GGGAAAGCTGTCTACTTAAGCTGACATTTGCTGGTGGAGTGAACACGCTGGCCCTGGCGCTGC	27266
QY	227	LeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnIleuArgLeuGlySerGln	246
Db	27267	CTGGAGGAATTCTGGGCCACATGGGGCCAGTTCCCTGGGGCCCCAGCTCCGCTCTGCCAA	27326
QY	247	ValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProTrp	266
Db	27327	GTGTACAGGGCTCTGGCGCTCGGGGAGGGTCTTCCCTCGAGATCCGACACCTCCCCAG	27386
QY	267	AlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis	284
Db	27387	ACCCACACTAAAGGCTGCCCTTCTCTACACATTTTGGACATTTCCAAAGTTCCAC	27440

RESULT	15
AC119115	
LOCUS	AC119115
DEFINITION	AC119115 138792 bp DNA linear HTG 18-JUN-2002
ACCESSION	Rattus norvegicus clone CH230-320N23, *** SEQUENCING IN PROGRESS
VERSION	***, 32 unordered pieces.
KEYWORDS	AC119115
SOURCE	AC119115.2 GI:21746718
ORGANISM	HTG: HTGS PHASE1.
	Norway rat.
	Rattus norvegicus
	Rattus norvegicus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
	Rattus.
REFERENCE	1 (bases 1 to 138792)
AUTHORS	Muzny,D.M., Adams,C., Adio-Oduola,B., All-oshman,F.R., Allen,C.,

REFERENCE  
AUTHORS  
1 (bases 1 to 138792)  
Muzny,D.M., Adams,C., Adjo-Ofunola,B., Ali-olsman,F.R., Allen,C., Alshrocks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T., Barbata,J., Benton,J., Bimarge,K., Blankenburg,K., Bonini,D., Bouck,J., Bowle,S., Brileva,M., Brown,M., Brown,M., Bryant,N.P., Bulmay,K., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carion,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P., Gabris,J., Gao,J., Garcia,A., Garner,T., Garza,N., Gili,R., Goriell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homsl,F.F., Howard,S., Huber,J., Huily,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,O., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Iichcharge,O., Ilen,C., Iliu,J., Iliu,M., Loulsaged,H., Lozad,R., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Mattindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkenko,S., Oguni,M., Okunonu,G., Orangunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., RojuboKan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Syatek,A., Tabor,P., Tamerisa,A., Tameris,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Umanai,K., Vasquez,L., Verra,Y., Villalon,D., Vinson,R., Wang,O., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S., Williams,G., Williamson,A., Wlecyxk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Unpublished  
Direct Submission  
2 (bases 1 to 138792)  
Worley,K.C.  
Direct Submission  
Submitted (25-APR-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 138792)  
Worley,K.C.  
Direct Submission  
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jul 14, 2002 this sequence version replaced gi:20303440.

----- Genome Center  
Center: Baylor College of Medicine  
Contig code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>  
 Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
 ----- Project Information -----  
 Center project name: GDXG  
 Center clone name: CH230-320N23  
 ----- Summary Statistics -----  
 Sequencing vector: Plasmid:  
 Chemistry: Dye-terminator Big Dye: 100% of reads  
 Assembly program: Phrap: version 0.990329  
 Consensus quality: 112/20 bases at least Q40  
 Consensus quality: 116/66 bases at least Q30  
 Consensus quality: 119/65 bases at least Q20  
 -----  
 \* NOTE: Estimated insert size may differ from sequence length  
 \* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 32 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1	1022:	contig of 1022 bp in length
1023	1122:	gap of unknown length
1123	2394:	contig of 1272 bp in length
2395	2494:	gap of unknown length
2495	3663:	contig of 1169 bp in length
3664	3763:	gap of unknown length
3764	4879:	contig of 1116 bp in length
4880	4979:	gap of unknown length
4980	6787:	contig of 1808 bp in length
6788	6887:	gap of unknown length
6888	8192:	contig of 1305 bp in length
8193	8292:	gap of unknown length
8293	9461:	contig of 1169 bp in length
9462	9561:	gap of unknown length
9562	11435:	contig of 1874 bp in length
11436	11535:	gap of unknown length
11536	13535:	contig of 2000 bp in length
13536	13635:	gap of unknown length
13636	17323:	contig of 3688 bp in length
17324	17423:	gap of unknown length
17424	20406:	contig of 2983 bp in length
20407	20506:	gap of unknown length
20507	22445:	contig of 2939 bp in length
23446	23545:	gap of unknown length
23546	25600:	contig of 2055 bp in length
25601	25700:	gap of unknown length
25701	28049:	contig of 2349 bp in length
28050	28149:	gap of unknown length
28150	30474:	contig of 2325 bp in length
30475	30574:	gap of unknown length
30575	35498:	contig of 4924 bp in length
35499	35598:	gap of unknown length
35599	39284:	contig of 3686 bp in length
39285	39384:	gap of unknown length
39385	41870:	contig of 2586 bp in length
41971	42070:	gap of unknown length
42071	45659:	contig of 3589 bp in length
45660	45759:	gap of unknown length
45760	50332:	contig of 4573 bp in length
50333	50432:	gap of unknown length
50433	54695:	contig of 4263 bp in length
54696	54795:	gap of unknown length
54796	58289:	contig of 3494 bp in length
58290	58389:	gap of unknown length
58390	63599:	contig of 5210 bp in length
63600	63699:	gap of unknown length
63700	70305:	contig of 6606 bp in length
70306	70405:	gap of unknown length
70406	76123:	contig of 5718 bp in length
76124	76223:	gap of unknown length

76224	84961:	contig of 8738 bp in length
84962	85061:	gap of unknown length
85062	93614:	contig of 8533 bp in length
93615	93714:	gap of unknown length
93715	103352:	contig of 9638 bp in length
103353	103452:	gap of unknown length
103453	110299:	contig of 6847 bp in length
110300	110399:	gap of unknown length
110400	116180:	contig of 5781 bp in length
116181	116280:	gap of unknown length
116281	125484:	contig of 9204 bp in length
125485	125584:	gap of unknown length
125585	138792:	contig of 13208 bp in length.

FEATURES  
 source 1.138792  
 /organism="Rattus norvegicus"  
 /db\_xref="taxon:10116"  
 /clone="CH230-320N23"

BASE COUNT 34391 a 31996 c 31712 g 36193 t 4500 others

ORIGIN

Alignment Scores:  
 Pred. No.: 1.7e-15 Length: 138792  
 Score: 489.50 Matches: 102  
 Percent Similarity: 64.50% Conservative: 7  
 Best Local Similarity: 60.36% Mismatches: 53  
 Query Match: 33.90% Indels: 53  
 DB: 2 Gaps: 1

US-09-245-198a-4 (1-284) x AC119115 (1-138792)

QY	160	GLYVALASPGLYTHYRVALSERGLYTRPGIUGLUALAARGILASNSERSESRPROLEU	179
DB	137345	GGTGGATGGAGACGTGGTGGTGGAGAGACCAATATACACGTCCAGCCCTCG	137404
QY	180	ARGTYRASNARGLNIIEGLYGLUPHEILEVALTHIRAGLALGLYLEUTYRILEUTYR	199
DB	137405	CGCTAGACGCCGACAGATTGGGGAATTTAGGTCATCAGAGGCGTGTCTACACCTGAC	137464
QY	200	CYS-----	200
DB	137465	TTTCAGATAGACCCCTGGCTCCATGGTAGACGATGGCTAAGGGAGGAGGCTTGACA	137524
QY	200	-----	200
DB	137525	AGGATGGTGGAGTGGAGAACCTGGCTTCATGACGAAGAGCCCTGATTTCGATG	137584
QY	201	-----	201
DB	137585	AGGGCAGCAGAGGTGTGATTTTGTGCTTCTCTCTCCTCCAGGTGACCTTGATGAG	137644
QY	207	GLYLSALVALTYRILEULYSLEULSPLLEULVALSPGLVALLEULALAEULARGCYS	226
DB	137645	GGGAAGGACGTACTGACCTGAGCTGCTGTAATGTGCTGCGCTGCGCTGC	137704
QY	227	LEUGLUGLPHSESRALAHRAALASERISLEULYPROGLINLEULARGLEUCYSGLN	246
DB	137705	CTGGAGAAFTTCTCACCCACACAGCAGCTCTCTGCGGCCCCAGCTCCGTGTGTCAG	137764
QY	247	VALISERGLYLEULALAEULARGPROGLYSERISLEULARGILEUTHYRLEUPROTP	266
DB	137765	GTGTGCGGCTGTGCTGTGCTGTGACACAGGCTCTCCCTTCGGAATCGATACCTTGC	137824
QY	267	ALAHLSLEULSALALAPROPHLEU 275	
DB	137825	ACTCATATTAAAGCGGCTACCTTCGTT 137851	

Search completed: May 8, 2003, 03:45:22  
 Job time : 2530.92 secs

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 02:07:02 ; Search time 36,2672 Seconds  
(without alignments)  
1043,455 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 1444

Sequence: 1 MSLLDFEISARLPLPRSLG.....PMAHKAAPFLTYRGLFQVH 284

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A.GeneSeq\_101002:\*

1: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT:\*  
2: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT:\*  
3: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT:\*  
4: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT:\*  
5: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT:\*  
6: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT:\*  
7: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT:\*  
8: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT:\*  
9: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT:\*  
10: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT:\*  
11: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT:\*  
12: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT:\*  
13: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT:\*  
14: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT:\*  
15: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT:\*  
16: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT:\*  
17: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT:\*  
18: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT:\*  
19: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT:\*  
20: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT:\*  
21: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT:\*  
22: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT:\*  
23: /SID2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1444	100.0	284	19	AAW47525 Homo sapiens tumour
2	1268	87.8	249	20	AAV09369 Human tumour necro
3	1268	87.8	249	21	AAW07526 Amino acid sequenc
4	1268	87.8	249	21	AAV95338 Human PRO207 antit
5	1268	87.8	249	23	AAU86129 Human PRO207 poly
6	1265	87.6	249	19	AAW29745 TNF related endoth
7	1265	87.6	249	22	AAE00891 Human TRPRA (TNF r
8	1066	73.8	273	22	AAU03499 TWAK extracellular
9	1062	73.5	208	20	AAW93590 Human TNRL3 protei
10	1020	70.6	225	19	AAW47524 Mus musculus tumou

11	1020	70.6	225	21	AAW07527	Amino acid sequenc
12	968	67.0	211	20	AAW93591	Mouse TNRL3 protei
13	792	54.8	189	19	AAW29746	TNF related endoth
14	792	54.8	189	22	AAE00892	Human UTAflag TRP
15	761	52.7	146	22	AAE00895	Human TRPRA (TNF r
16	716	50.0	325	22	ABB67553	Drosophila melanog
17	116	8.0	409	23	AAU77718	Drosophila melanog
18	108.5	7.5	211	21	AAV58216	Canine mature CD15
19	108.5	7.5	260	21	AAV58215	Canine CD154, Can
20	106.5	7.4	254	16	AAW64190	Human 4-1BB-L poly
21	106.5	7.4	254	18	AAW26657	Human 4-1BB ligand
22	106.5	7.4	254	23	ABB75953	Human cytokine 4-1
23	106	7.3	1428	21	AAV97033	Caspase 8-interact
24	105.5	7.3	406	23	AAU77717	Drosophila melanog
25	104.5	7.2	779	23	ABB07845	Human MAP kinase p
26	104	7.2	409	23	AAU77716	Drosophila melanog
27	104	7.2	1323	15	AAW55248	N-methyl-D-aspart
28	102	7.1	256	22	AAW25657	Human protein sequ
29	100.5	7.0	647	17	AAW04327	Rat peltin. Ratu
30	100	6.9	220	22	AAW62340	Gp120 V3 loop-CD15
31	99	6.9	574	21	AAV97032	Caspase 8-interact
32	99	6.9	1008	22	AAW78891	Human protein Sg
33	99	6.9	1020	22	AAW79875	Human protein Sg
34	96.5	6.7	1097	22	ABG25655	Novel human diagno
35	96.5	6.7	1631	22	ABG22481	Novel human diagno
36	96	6.6	234	22	AAW62339	Gp120 V3 loop-CD15
37	95	6.6	240	23	AAE13680	Human HYEM-binding
38	95	6.6	294	13	AAW69956	NF-kB receptor act
39	95	6.6	294	19	AAW68292	NF-kB receptor act
40	95	6.6	294	19	AAW68292	Murine receptor ac
41	95	6.6	294	22	AAE08737	Murine RANKL (rece
42	95	6.6	294	22	AAE04425	Murine RANKL (rece
43	94.5	6.5	1006	22	AAE01992	Novel human diagno
44	94	6.5	876	22	ABG00217	Novel human diagno
45	93.5	6.5	409	22	AAW00076	Human polypeptide

#### ALIGNMENTS

RESULT 1	
AAW47525	AAW47525 standard; Protein; 284 AA.
XX	
XX	
AC	AAW47525;
XX	
DT	21-JUL-1998 (first entry)
XX	
DE	Homo sapiens tumour necrosis factor related ligand (TNRL).
XX	
KW	TNRL; tumour necrosis factor related ligand; tnfr; treatment;
KW	cancer; autoimmune disease; immune system; stimulation; suppression;
KW	graft rejection;
XX	
OS	Homo sapiens.
XX	
PN	WO9805783-A1.
XX	
PD	12-FEB-1998.
XX	
PF	07-AUG-1997; 97WO-US13945.
XX	
PR	18-MAR-1997; 97US-0040820.
PR	07-AUG-1996; 96US-0023541.
XX	
PR	18-OCT-1996; 96US-0028515.
PA	(BIOJ) BIOGEN INC.
PA	(UYGE) UNIV GENEVA FACULTY MEDICINE.
XX	
PI	Browning JL, Chicheportiche Y;
XX	
DR	WPI, 1998-145619/13.
DR	N-PSDB; AAV18600.

XX Tumour necrosis factor related ligand - useful for, e.g. treating  
 PT cancer, auto-immune disease and immune responses to tissue grafts  
 XX  
 PS Claim 12; Pages 50-51; 69pp; English.

CC The sequence is that of human tumour necrosis factor related  
 CC ligand (TRELL). TRELL or active fragments can be included with a  
 CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
 CC diseases or immune responses to tissue grafts, or to stimulate or  
 CC suppress the immune system. It is useful to screen for TRELL  
 CC receptors, by labelling with a detectable label and screening  
 CC compositions for binding. Agents interfering with TRELL-receptor  
 CC binding can also be screened for, can then be administered,  
 CC optionally with interferon- gamma, to induce cell death or  
 CC treat, suppress or alter immune responses (especially involving human  
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
 CC receptor. It's coding sequence can be used in gene therapy for  
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
 CC autoimmune and inflammatory diseases or inherited genetic disorders,  
 CC by introducing into cells, and expressing, therapeutically effective  
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
 CC It may also be of use in the preparation of probe probes for  
 CC screening natural/synthetic DNAs for TRELL-encoding sequences  
 CC and for antisense therapy.

XX Sequence 284 AA;

Query Match 100.0%; Score 1444; DB 19; Length 284;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-129;  
 Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSLLDFEISARRLPRISGSDGAVROAPAPMAARRRRGRGCEPTALLVPLA 60  
 DB 1 MSLLDFEISARRLPRISGSDGAVROAPAPMAARRRRGRGCEPTALLVPLA 60  
 OY 61 LGLGIALACGLLLAVNSLSGRASLSAQEPAGEELVAEEDDPSLEINPQTESODPAPFL 120  
 DB 61 LGLGIALACGLLLAVNSLSGRASLSAQEPAGEELVAEEDDPSLEINPQTESODPAPFL 120  
 OY 121 NRVLPBRASAPKGRKTRARRAIAAHYEVHPRPGOGAAGVDGTVSGMEARINSSPLR 180  
 DB 121 NRVLPBRASAPKGRKTRARRAIAAHYEVHPRPGOGAAGVDGTVSGMEARINSSPLR 180  
 OY 181 YNRQIGERTYRAGLYLYCOVHFDEGKAVYLKLDLVDGYLALRCLFEFSATASSIGP 240  
 DB 181 YNRQIGERTYRAGLYLYCOVHFDEGKAVYLKLDLVDGYLALRCLFEFSATASSIGP 240  
 OY 241 QLRLCOVSGLLALRPGSSLRIRTLPMWHLKAPFLTYGFLQVH 284  
 DB 241 QLRLCOVSGLLALRPGSSLRIRTLPMWHLKAPFLTYGFLQVH 284

RESULT 2

AA09369  
 ID AA09369 standard; protein: 249 AA.

AC AA09369;

DT 15-JUL-1999 (first entry)

DE Human tumour necrosis factor Apo-3 ligand protein sequence.

KW Human: tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;

KM NF-kappaB-dependent transcription; JNK/SAPK-dependent response;

XX cancer.

OS Homo sapiens.

XX MO9919490-A1.

PD 22-APR-1999.

PF 09-OCT-1998; 98WO-US21407.  
 XX  
 PR 17-DEC-1997; 97US-0069862.  
 XX 10-OCT-1997; 97US-0062037.  
 PR

PA (GERTH) GENENTECH INC.

PI Ashkenazi AJ, Marsters SA, Pittl R;

DR WPI: 1999-287982/24.

DR N-PSDB; AAX56000.

PT New human Apo-3 ligand (a tumour necrosis factor) homologue

PS Claim 1; Fig 1; 74pp; English.

CC The present sequence represents a human tumour necrosis factor (TNF)  
 CC and lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has  
 CC cytostatic activity. Apo-3 ligand can be used to induce apoptosis in  
 CC mammalian cancer cells, to induce NF-kappaB-dependent transcription and  
 CC to induce JNK/SAPK-dependent responses in mammalian cells.

XX Sequence 249 AA;

Query Match 87.8%; Score 1268; DB 20; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-113;  
 Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 MAARRSRRRRGRGREGPALLVPLALGILACIGLLAVNSLSGRASLSAQEPAGEEL 95  
 DB 1 MAARRSRRRRGRGREGPALLVPLALGILACIGLLAVNSLSGRASLSAQEPAGEEL 95  
 OY 96 VAEEDQDPSLEINPQTESODPAPFLNRLVRRRSAPKGRKTRARRAIAAHYEVHPRGOD 155  
 DB 61 VAEEDQDPSLEINPQTESODPAPFLNRLVRRRSAPKGRKTRARRAIAAHYEVHPRGOD 120  
 OY 156 GAQAGVDGTVSGMEARINSSPLRYNROIGEFYTRAGLYLYCOVHFDEGKAVYLKLD 215  
 DB 121 GAQAGVDGTVSGMEARINSSPLRYNROIGEFYTRAGLYLYCOVHFDEGKAVYLKLD 180  
 OY 216 LLDVGVIALRCLFEFSATASSLSGOLRLCOVSGLLALRPGSSLRIRTLPMWHLKAPFL 275  
 DB 181 LLDVGVIALRCLFEFSATASSLSGOLRLCOVSGLLALRPGSSLRIRTLPMWHLKAPFL 240  
 OY 276 TYFGFLQVH 284  
 DB 241 TYFGFLQVH 249

RESULT 3

AAB07526  
 ID AAB07526 standard; protein: 249 AA.

AC AAB07526;

DT 20-OCT-2000 (first entry)

DE Amino acid sequence of a soluble recombinant human TWEAK protein.

KW TWEAK protein; immunological disorder; immune response; inflammation;

KM TWEAK blocking agent; autoimmune disease; organ transplant rejection;

KW Graft-versus-host disease; GVHD; lymphoid cell malignancy; shock; tumour.

XX Homo sapiens.

XX WO200042073-A1.

PD 20-JUL-2000.

PF 14-JAN-2000; 2000WO-US01044.

PR 15-JAN-1999; 99US-0116168.

PA (BIO ) BIOGEN INC.  
 XX  
 PI Remmert P;  
 XX  
 DR WPI; 2000-476036/41.  
 XX  
 PT Preventing and treating immune responses using modulators, especially  
 PT antibodies of TWEAK. TWEAK receptors and TWEAK ligands, useful for  
 PT treating e.g. inflammation and graft versus host disease -  
 XX  
 PS Disclosure; Fig 1; 45pp; English.  
 XX  
 CC The present sequence represents a TWEAK protein. The specification  
 CC describes a method for preventing or treating an immunological  
 CC disorder and/or inhibiting an immune response in an animal. The  
 CC method comprises administering a TWEAK blocking agent. The method may  
 CC be used for preventing and treating immune disorders associated with  
 CC inappropriate expression and/or activity of TWEAK. These disorders  
 CC include autoimmune diseases, acute and chronic inflammation, organ  
 CC transplant rejection, graft-versus-host disease (GVHD), lymphoid cell  
 CC malignancies, septic and other forms of shock, loss of immune  
 CC responsiveness (as seen in human immunodeficiency virus (HIV)  
 CC infections) and failure of the immune response to tumour growth.  
 CC  
 SQ Sequence 249 AA;  
 Query Match 87.8%; Score 1268; DB 21; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 6,6e-113;  
 Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 MAARRSQRRRGRRGEPGTALVPLALGLALGLALGLALLAVSLGSRASLSAEPAGEEL 95  
 DB 1 MAARRSQRRRGRRGEPGTALVPLALGLALGLALGLALLAVSLGSRASLSAEPAGEEL 60  
 QY 96 VAEEDDPSLNPQTEESODPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEYHPRGOD 155  
 DB 61 VAEEDDPSLNPQTEESODPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEYHPRGOD 120  
 QY 156 GAAGAGDGTYSGWGEARINSSPLRYNRQIGETIVTRAGIYLYCOVHPEBGAAVYLKLD 215  
 DB 121 GAAGAGDGTYSGWGEARINSSPLRYNRQIGETIVTRAGIYLYCOVHPEBGAAVYLKLD 180  
 QY 216 LTVDGVALRLCEEFSAATASSGLPOLRLCOVSGLLALRPGSSLRITPLMAHLKAPFL 275  
 DB 181 LTVDGVALRLCEEFSAATASSGLPOLRLCOVSGLLALRPGSSLRITPLMAHLKAPFL 240  
 QY 276 TYFGLEFQVH 284  
 DB 241 TYFGLEFQVH 249

RESULT 4  
 AAY95338  
 ID AAY95338 standard; Protein; 249 AA.  
 XX  
 AC AAY95338;  
 XX  
 DT 25-SEP-2000 (first entry)  
 XX  
 DE Human PRO207 antitumour protein.  
 XX  
 KW PRO207; human; antitumour; tumour; therapy; cytostatic;  
 KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;  
 KW uterine cancer; prostate cancer; lung cancer; bladder cancer;  
 KW central nervous system cancer; melanoma; leukaemia; neoplasm.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Reptide 1..40  
 FT Protein /label= Signal\_peptide 41..249  
 FT /label= PRO207

FT Modified-site 27..33  
 FT /note= "N-myristoylation"  
 FT Modified-site 29..35  
 FT /note= "N-myristoylation"  
 FT Modified-site 36..42  
 FT /note= "N-myristoylation"  
 FT Modified-site 45..51  
 FT /note= "N-myristoylation"  
 FT Modified-site 118..124  
 FT /note= "N-myristoylation"  
 FT Modified-site 121..127  
 FT /note= "N-myristoylation"  
 FT Modified-site 125..131  
 FT /note= "N-myristoylation"  
 FT Modified-site 128..134  
 FT /note= "N-myristoylation"  
 FT Modified-site 139..143  
 FT /note= "N-myristoylation"  
 FT Modified-site 140..14  
 FT /note= "Asn is N-glycosylated"  
 FT Modified-site 97..101  
 FT /note= "amidation"  
 FT Peptide 24..35  
 FT /note= "prokaryotic membrane lipid protein lipid"  
 XX  
 PN MO200037638-A2.  
 XX  
 PD 29-JUN-2000.  
 XX  
 PF 02-DEC-1999; 99WO-US28565.  
 XX  
 PR 22-DEC-1998; 98US-0113296.  
 PR 08-MAR-1999; 99WO-US05028.  
 PR 21-APR-1999; 99US-0130332.  
 PR 28-APR-1999; 99US-0131445.  
 PR 14-MAY-1999; 99US-0134287.  
 PR 20-JUL-1999; 99US-0144758.  
 PR 26-JUL-1999; 99US-0145698.  
 PR 15-SEP-1999; 99WO-US21090.  
 PR 15-SEP-1999; 99WO-US21547.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;  
 PI Napier MA, Pitti RM, Wood WI;  
 XX  
 DR WPI; 2000-442668/38.  
 XX  
 DR N-PDB; AAA49717.  
 XX  
 PT Novel composition to inhibit neoplastic cell growth or for treating  
 PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,  
 PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO356, PRO509 or  
 PT PRO866  
 XX  
 PS Claim 19; Fig 4; 172pp; English.  
 XX  
 CC The present sequence is that of human antitumour protein PRO207,  
 CC as deduced from a foetal kidney cDNA clone (see AAA49717). PRO207  
 CC shows amino acid sequence identity to tumour necrosis factor family  
 CC members, especially human lymphotoxin-beta (23.4%) and human CD40  
 CC ligand (19.8%). Mol. wt. is 27,216. A claimed method for inhibiting  
 CC the growth of a tumour cell comprises exposing the tumor cell  
 CC to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224, PRO328, PRO301,  
 CC PRO526, PRO362, PRO509 or PRO866 (see AAY95337-49), their  
 CC agonists or chimeric polypeptides incorporating them. The tumour  
 CC is especially a cancer selected from breast, ovarian, renal,  
 CC colorectal, uterine, prostate, lung, bladder and central nervous  
 CC system cancer, melanoma and leukaemia. Methods for the recombinant  
 CC expression of the antitumour proteins are also provided.  
 XX  
 SQ Sequence 249 AA;  
 Query Match 87.8%; Score 1268; DB 21; Length 249;

Best Local Similarity 100.0%; Pred. No. 6.6e-113;  
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 MAARRQRRRRGRGEGTALLVPLALGLALACGLLAVVSLGSRASLSAQEPAGEEL 95  
DB 1 MAARRQRRRRGRGEGTALLVPLALGLALACGLLAVVSLGSRASLSAQEPAGEEL 60  
QY 96 VAEEDDPSELNPQTEESODPAPFLNRLVPRRSAPKGRKTRARRAIAHYEVHPRPGD 155  
DB 61 VAEEDDPSELNPQTEESODPAPFLNRLVPRRSAPKGRKTRARRAIAHYEVHPRPGD 120  
QY 156 GAQAGVDGTVSGMEARINSSPLRYNRQIGEFITVRAGLYLYLCQVHDEGKAVYLLKD 215  
DB 121 GAQAGVDGTVSGMEARINSSPLRYNRQIGEFITVRAGLYLYLCQVHDEGKAVYLLKD 180  
QY 216 LTVDGVALRCLEEFSSATTAASSLGPOLRLCOVSGLLALRPGSSLRIRTLPMHNLKAAPFL 275  
DB 181 LTVDGVALRCLEEFSSATTAASSLGPOLRLCOVSGLLALRPGSSLRIRTLPMHNLKAAPFL 240  
QY 276 TYFGLEFOVH 284  
DB 241 TYFGLEFOVH 249

## RESULT 5

AAU86129  
ID AAU86129 standard; Protein; 249 AA.

AC AAU86129;  
DT 15-JUL-2002 (first entry)  
DE Human PRO207 polypeptide.  
KW Human: PRO; benign tumour; malignant tumour; lymphoid malignancy;  
KW leukaemia; neuronal disorder; stromal disorder; blastocoeal disorder;  
KW inflammatory disorder; immune disorder; angiogenic disorder;  
KW cytostatic; neuroprotective.  
OS Homo sapiens.  
PN WO200153486-A1.  
PD 26-JUL-2001.

XX 11-FEB-2000; 2000WO-US03565.  
XX 08-MAR-1999; 99WO-US05028.  
XX 11-MAR-1999; 99US-123972P.  
XX 11-MAY-1999; 99US-133459P.  
XX 02-JUN-1999; 99WO-US12252.  
XX 22-JUN-1999; 99US-140650P.  
XX 22-JUN-1999; 99US-140653P.  
XX 20-JUL-1999; 99US-144758P.  
XX 26-JUL-1999; 99US-145698P.  
XX 28-JUL-1999; 99US-146222P.  
XX 17-AUG-1999; 99US-149395P.  
XX 31-AUG-1999; 99US-151689P.  
XX 01-SEP-1999; 99WO-US20111.  
XX 15-SEP-1999; 99WO-US21090.  
XX 30-NOV-1999; 99WO-US28313.  
XX 01-DEC-1999; 99WO-US28301.  
XX 01-DEC-1999; 99WO-US28634.  
XX 05-JAN-2000; 2000WO-US00219.

XX (SETH ) GENENTECH INC.

XX Ashkenazi AV, Goddard A, Godowski PJ, Gurney AL, Hillan KT,  
PI Marsters SA, Pan J, Pittl RM, Roy MA, Smith V, Stone DM;  
XX Matanabe CK, Wood WI;

DR WPI; 2002-205567/26.  
DR N-PSDB; ABR40255.

XX Thirty five nucleic acids encoding PRO polypeptides, useful for  
PT treating benign or malignant tumours, leukaemias and lymphoid  
PT malignancies, inflammatory, angiogenic and immunologic disorders -  
XX  
PS Claim 61: Fig 4; 302pp; English.

XX The present invention relates to the isolation of novel human PRO  
CC polypeptides and the polynucleotide sequences encoding them. The  
CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are  
CC useful for treating benign or malignant tumours (e.g. renal, kidney,  
CC bladder, breast, etc), leukaemias and lymphoid malignancies, other  
CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,  
CC macrophagal, stromal and blastocoeal disorders, inflammatory, immune  
CC and angiogenic disorders. The polynucleotide sequences are also  
CC useful in gene therapy. AAU86128-AAU86162 represent the human PRO  
CC polypeptides of the invention.

SQ Sequence 249 AA:

Query Match 87.8%; Score 1268; DB 23; Length 249;  
Best Local Similarity 100.0%; Pred. No. 6.6e-113;  
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 MAARRQRRRRGRGEGTALLVPLALGLALACGLLAVVSLGSRASLSAQEPAGEEL 95  
DB 1 MAARRQRRRRGRGEGTALLVPLALGLALACGLLAVVSLGSRASLSAQEPAGEEL 60  
QY 96 VAEEDDPSELNPQTEESODPAPFLNRLVPRRSAPKGRKTRARRAIAHYEVHPRPGD 155  
DB 61 VAEEDDPSELNPQTEESODPAPFLNRLVPRRSAPKGRKTRARRAIAHYEVHPRPGD 120  
QY 156 GAQAGVDGTVSGMEARINSSPLRYNRQIGEFITVRAGLYLYLCQVHDEGKAVYLLKD 215  
DB 121 GAQAGVDGTVSGMEARINSSPLRYNRQIGEFITVRAGLYLYLCQVHDEGKAVYLLKD 180  
QY 216 LTVDGVALRCLEEFSSATTAASSLGPOLRLCOVSGLLALRPGSSLRIRTLPMHNLKAAPFL 275  
DB 181 LTVDGVALRCLEEFSSATTAASSLGPOLRLCOVSGLLALRPGSSLRIRTLPMHNLKAAPFL 240  
QY 276 TYFGLEFOVH 284  
DB 241 TYFGLEFOVH 249

## RESULT 6

AAW29745  
ID AAW29745 standard; Protein; 249 AA.

XX AAW29745;  
DT 27-OCT-1998 (first entry)  
DE TNF related endothelium proliferative agent protein.  
KW TNF; endothelium proliferative agent; TREPA; wound healing; cancer;  
KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.  
OS Homo sapiens.

PN WO9835061-A2.  
PD 13-AUG-1998.

XX 12-FEB-1998; 98WO-US02859.  
XX 10-FEB-1998; 98US-0021706.  
XX 12-FEB-1997; 97US-0798692.

XX (ABBO ) ABBOTT LAB.

XX WALLEY SR;



DR WPI: 1998-447255/38.  
 DR N-PSDB: AAV47613.  
 XX  
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 treatment of autoimmune disease; tumours and inflammation  
 XX  
 PS Claim 16; Page 123-4; 142pp; English.  
 XX  
 CC The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).  
 CC  
 SQ Sequence 249 AA;  
 XX  
 Query Match 87.6%; Score 1265; DB 19; Length 249;  
 Best Local Similarity 99.6%; Pred. No. 1.3e-112;  
 Matches 248; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 36 MAARRSORRRGRGEGTALLVPLALGLALACGLALLAVSLGSRASLSAEPQOEL 95  
 DB 1 MAARRSQKRRGRGEGTALLVPLALGLALACGLALLAVSLGSRASLSAEPQOEL 60  
 QY 96 VAEDDDPSELNPQTESODPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRGDD 155  
 DB 61 VAEDDDPSELNPQTESODPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRGDD 120  
 QY 156 GAAGVDTGTVSGWEAKINSSPLRYNRQIGETIVRAGLYLYXCQVHEDEGKAVYLKLD 215  
 DB 121 GAAGVDTGTVSGWEAKINSSPLRYNRQIGETIVRAGLYLYXCQVHEDEGKAVYLKLD 180  
 QY 216 LTVDGVALRCLKEFSATASSIGPOLRLCOVSGLLALRPGSSLRITLPMALIKKAPFL 275  
 DB 181 LTVDGVALRCLKEFSATASSIGPOLRLCOVSGLLALRPGSSLRITLPMALIKKAPFL 240  
 QY 276 TYRGLFOVH 284  
 DB 241 TYRGLFOVH 249  
 XX  
 RESULT 7  
 AAE00891  
 ID AAE00891 standard; Protein: 249 AA.  
 XX  
 AC AAE00891;  
 XX  
 DT 04-JUL-2001 (first entry)  
 XX  
 DE Human TREPA (TNF related endothelium proliferative agent).  
 XX  
 KM Human: tumour necrosis factor; TNF; angiogenesis; wound healing;  
 KM TREPA: TNF related endothelium proliferative agent; tumour; metastasis;  
 KM grafting; vulneryary.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 98..249  
 FT /label= Extracellular\_domain  
 XX  
 PN US6207642-B1.

PD 27-MAR-2001.  
 XX  
 PF 26-JUN-1998; 98US-0105343.  
 XX  
 PR 12-FEB-1997; 97US-0798692.  
 PR 10-FEB-1998; 98US-0021706.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Wiley SR;  
 XX  
 DR WPI: 2001-280760/29.  
 DR N-PSDB: AAD04350.  
 XX  
 PT Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 XX  
 PS Claim 1; Column 75-76; 53pp; English.  
 XX  
 CC The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts.  
 CC The present amino acid sequence is clone ID #690050 human TREPA.  
 CC  
 SQ Sequence 249 AA;  
 XX  
 Query Match 87.6%; Score 1265; DB 22; Length 249;  
 Best Local Similarity 99.6%; Pred. No. 1.3e-112;  
 Matches 248; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 XX  
 QY 36 MAARRSORRRGRGEGTALLVPLALGLALACGLALLAVSLGSRASLSAEPQOEL 95  
 DB 1 MAARRSQKRRGRGEGTALLVPLALGLALACGLALLAVSLGSRASLSAEPQOEL 60  
 QY 96 VAEDDDPSELNPQTESODPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRGDD 155  
 DB 61 VAEDDDPSELNPQTESODPAPFLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRGDD 120  
 QY 156 GAAGVDTGTVSGWEAKINSSPLRYNRQIGETIVRAGLYLYXCQVHEDEGKAVYLKLD 215  
 DB 121 GAAGVDTGTVSGWEAKINSSPLRYNRQIGETIVRAGLYLYXCQVHEDEGKAVYLKLD 180  
 QY 216 LTVDGVALRCLKEFSATASSIGPOLRLCOVSGLLALRPGSSLRITLPMALIKKAPFL 275  
 DB 181 LTVDGVALRCLKEFSATASSIGPOLRLCOVSGLLALRPGSSLRITLPMALIKKAPFL 240  
 QY 276 TYRGLFOVH 284  
 DB 241 TYRGLFOVH 249  
 XX  
 RESULT 8  
 AAU03499  
 ID AAU03499 standard; Protein: 273 AA.  
 XX  
 AC AAU03499;  
 XX  
 DT 26-SEP-2001 (first entry)  
 XX  
 DE TWEAK extracellular domain-containing fusion protein.  
 XX  
 KM TWEAK extracellular domain: tumour necrosis factor; TNF; angiogenesis;  
 KM ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 KM retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 KM rubecosis; uveitis; macular degeneration; arthritis; rheumatism;  
 KM corneal graft neovascularisation; psoriasis; metastatic condition;  
 KM malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;

preneoplastic condition; myocardial angiogenesis; wound granulation;  
 KM scleroderma; vascular adhesion; telangiectasia; ischaemia; human;  
 KM atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 KM peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;  
 KM fusion protein.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN WO200145730-A2.  
 PD 28-JUN-2001.  
 XX  
 PF 19-DEC-2000; 2000WO-US34755.  
 XX  
 PR 20-DEC-1999; 99US-0172878.  
 PR 10-MAY-2000; 2000US-0203347.  
 XX  
 PA (IMM) IMMUNEX CORP.  
 XX  
 PI Wiley SR;  
 PI  
 DR WPI: 2001-417975/44.  
 DR N-PSDB; AAS03964.  
 XX  
 PT Modulating angiogenesis in a mammal for treating diseases mediated by  
 PT angiogenesis e.g. solid tumours and vascular deficiencies of cardiac or  
 PT peripheral tissue, by administering antagonist or agonist of TWEAK  
 PT receptor.  
 XX  
 PS Example 1; Page 41; 46pp; English.  
 XX  
 CC The sequence represents a fusion protein encoded by the  
 CC expression vector pDC409-LZ-TWEAK. The fusion protein comprises a  
 CC growth hormone leader, a leucine zipper multimerisation domain, and  
 CC the extracellular domain of human TWEAK. The fusion protein was  
 CC used in the isolation of human TWEAK receptor (TWEAKR)-expressing  
 CC clones from a COS cell human cDNA library. The TWEAKR protein is  
 CC a member of the tumour necrosis factor (TNF) family and induces  
 CC angiogenesis. TWEAKR may therefore be used to screen for and  
 CC develop TWEAKR agonists and antagonists for the modulation of  
 CC angiogenesis, to be used in the treatment and diagnosis of human disease.  
 CC The disorders mediated by angiogenesis include ocular disorders  
 CC characterised by ocular neovascularisation such as diabetic retinopathy,  
 CC neovascular glaucoma, retinoblastoma, retinopathy of prematurity,  
 CC retrolental fibroplasia, rubeosis, uveitis, macular degeneration and  
 CC corneal graft neovascularisation, and inflammatory diseases such as  
 CC arthritis, rheumatism and psoriasis. Other treatable diseases include  
 CC malignant and metastatic conditions such as sarcomas and carcinomas,  
 CC benign tumours and preneoplastic conditions, myocardial angiogenesis,  
 CC haemophilic joints, scleroderma, vascular adhesions, atherosclerotic  
 CC plaque neovascularisation, telangiectasia, wound granulation, coronary  
 CC atherosclerosis, peripheral atherosclerosis and ischaemia.  
 XX  
 SQ Sequence 273 AA;  
 Query Match 73.8%; Score 1066; DB 22; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-93;  
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

78 SLGSRASLSAQEPAGEELVAEEDDPSELNPTQESODPAPFLNLTVPKRSAPKGRTR 137  
 DB SLGSRASLSAQEPAGEELVAEEDDPSELNPTQESODPAPFLNLTVPKRSAPKGRTR 126  
 QY 138 ARRAIAAHYEVHPRGQDGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYY 197  
 DB 127 ARRAIAAHYEVHPRGQDGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYY 186  
 QY 198 LYCOVHDEGKAVYIKLDLVGVIALRCLFEFSSTAASSIGPQRLCQVSGLLARPGS 257  
 DB 187 LYCOVHDEGKAVYIKLDLVGVIALRCLFEFSSTAASSIGPQRLCQVSGLLARPGS 246  
 QY 258 SLRIRTLPMWHLKAPFLTYFGLFOVH 284

DB 247 SLRIRTLPMWHLKAPFLTYFGLFOVH 273  
 RESULT 9  
 AAW93590  
 ID AAW93590 standard; Protein; 208 AA.  
 XX  
 AC AAW93590;  
 XX  
 DT 18-JUN-1999 (first entry)  
 XX  
 DE Human TNFR3 protein.  
 XX  
 KM Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;  
 KM developmental abnormality; gestational abnormality; prostate cancer;  
 KM APO6; APO8; APO9; TNFR-1; TNFR-3; diagnosis; treatment; therapy; disease;  
 KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KM apoptosis; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9911791-A2.  
 XX  
 PD 11-MAR-1999.  
 XX  
 PF 04-SEP-1998; 98WO-US18393.  
 XX  
 PR 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNIT) UNIT WASHINGTON.  
 XX  
 PI Chaudhary PM;  
 PI  
 DR WPI: 1999-205191/17.  
 DR N-PSDB; AAX23424.  
 XX  
 PT New Tumour Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 PT  
 PS Claim 40; Fig 13A; 156pp; English.  
 XX  
 CC This invention describes isolated Tumour Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNFR1 and TNFR3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC moiety. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the change in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active  
 CC fragments and APO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
 CC activity. The method is performed in vivo or in vitro. APO polypeptides  
 CC are all useful as immunogens for preparing antibodies. APO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 XX  
 SQ Sequence 208 AA;  
 Query Match 73.5%; Score 1062; DB 20; Length 208;  
 Best Local Similarity 99.5%; Pred. No. 2.6e-93;  
 Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

77 VSLSRASLSAQEPAGEELVAEEDDPSELNPTQESODPAPFLNLTVPKRSAPKGRTR 136  
 VSLSRASLSAQEPAGEELVAEEDDPSELNPTQESODPAPFLNLTVPKRSAPKGRTR 136

Db 1 VSLGSRASLSAQEPAPQELVAEEDDPSELNPQTEESQDPAPFLNRLVARRSAPGRKT 60  
 QY 137 RARRAIAAHYEHVPRPGDGAQAGVDGYVSGWEARINSSPLRTNRQIGETIVTRAGLY 196  
 Db 61 RARRAIAAHYEHVPRPGDGAQAGVDGYVSGWEARINSSPLRTNRQIGETIVTRAGLY 120  
 QY 197 YLYCOVHFDEGKAVYTKLDLVGVYALALRCLFEFSATASSISGPOLRLCQVSGLLALRPG 256  
 Db 121 YLYCOVHFDEGKAVYTKLDLVGVYALALRCLFEFSATASSISGPOLRLCQVSGLLALRPG 180  
 QY 257 SLSLRITLPMWHLKAKAPFLTYFGLFQVH 284  
 Db 181 SLSLRITLPMWHLKAKAPFLTYFGLFQVH 208

RESULT 10  
 AAB07524  
 ID AAB07524 standard; Protein: 225 AA.  
 AC AAB07524;  
 XX  
 DT 21-JUL-1998 (first entry)  
 XX  
 DE Mus musculus tumour necrosis factor related ligand (TRELL).  
 XX  
 KW TRELL: tumour necrosis factor related ligand; tnf; treatment:  
 KW cancer; autoimmune disease; immune system; stimulation; suppression;  
 KM graft rejection.  
 XX  
 OS Mus musculus.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 1..21  
 FT /note= "hydrophobic, transmembrane domain"  
 XX  
 PN WO9605783-A1.  
 XX  
 DE 12-FEB-1998.  
 PD  
 PE 07-AUG-1997; 97WO-US13945.  
 XX  
 PR 18-MAR-1997; 97US-0040820.  
 PR 07-AUG-1996; 96US-0023541.  
 PR 18-OCT-1996; 96US-0028515.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 PA (UYGE-) UNITIV GENEVA FACULTY MEDICINE.  
 XX  
 PI Browning JL, Chicheportliche Y;  
 XX  
 DR WPI: 1998-145619/13.  
 DR N-PSDB; AAV18599.  
 XX  
 PT Tumour necrosis factor related ligand - useful for, e.g. treating  
 PT cancer, auto-immune disease and immune responses to tissue grafts  
 XX  
 PS Claim 12; Pages 48-50; 69pp; English.

The sequence is that of mouse tumour necrosis factor related ligand (TRELL). TRELL or active fragments can be included with a carrier in pharmaceutical compositions to treat cancer, autoimmune diseases or immune responses to tissue grafts, or to stimulate or suppress the immune system. It is useful to screen for TRELL receptors, by labelling with a detectable label and screening compositions for binding. Agents interfering with TRELL-receptor binding can also be screened for, can then be administered, optionally with interferon- gamma, to induce cell death or treat, suppress or alter immune responses (especially involving human adenocarcinoma cells) involving a signal pathway between TRELL and its receptor. It's coding sequence can be used in gene therapy for TRELL-related disorders in mammals (especially humans), e.g. tumours, autoimmune and inflammatory diseases or inherited genetic disorders, by introducing into cells, and expressing, therapeutically effective

CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
 CC It may also be of use in the preparation of probe probes for  
 CC screening natural/synthetic DNAs for TRELL-encoding sequences  
 CC and for antisense therapy.  
 XX  
 SQ Sequence 225 AA;  
 Query Match 70.6%; Score 1020; DB 19; Length 225;  
 Best Local Similarity 88.8%; Pred. No. 2,9e-89;  
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY 61 LSLGLALACIGLLAVVSLGSRASLSAQEPAPQELVAEEDDPSELNPQTEESQDPAPFL 120  
 Db 2 LSLGLALACIGLLAVVSLGSRASLSAQEPAPQELVAEEDDPSELNPQTEESQDPAPFL 61  
 QY 121 NRVPRRSAPGRKTRARRAIAAHYEHVPRPGDGAQAGVDGYVSGWEARINSSPLR 180  
 Db 62 EQLVPRRSAPGRKTRARRAIAAHYEHVPRPGDGAQAGVDGYVSGWEARINSSPLR 121  
 QY 181 YNRQIGETIVTRAGLYYLYCOVHFDEGKAVYTKLDLVGVYALALRCLFEFSATASSISGP 240  
 Db 122 YNRQIGETIVTRAGLYYLYCOVHFDEGKAVYTKLDLVGVYALALRCLFEFSATASSISGP 181  
 QY 241 QLRLCQVSGLLALRPGSSLSLRITLPMWHLKAKAPFLTYFGLFQVH 284  
 Db 182 QLRLCQVSGLLALRPGSSLSLRITLPMWHLKAKAPFLTYFGLFQVH 225

RESULT 11  
 AAB07527  
 ID AAB07527 standard; protein: 225 AA.  
 AC AAB07527;  
 XX  
 DT 20-OCT-2000 (first entry)  
 XX  
 DE Amino acid sequence of a soluble recombinant murine TWEAK protein.  
 XX  
 KW TWEAK protein; immunological disorder; immune response; inflammation;  
 KW TWEAK blocking agent; autoimmune disease; organ transplant rejection;  
 KW Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock; tumour.  
 XX  
 OS Mus sp.  
 XX  
 PN WO200042073-A1.  
 XX  
 PD 20-JUL-2000.  
 XX  
 PE 14-JAN-2000; 2000WO-US01044.  
 XX  
 PR 15-JAN-1999; 99US-0116168.  
 XX  
 PA (BIOJ ) BIOGEN INC.  
 PA Rennert P;  
 PI  
 XX  
 DR WPI: 2000-476036/41.  
 XX  
 PT Preventing and treating immune responses using modulators, especially  
 PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for  
 PT treating e.g. inflammation and graft versus host disease -  
 XX  
 PS Disclosure; Fig 1; 45pp; English.

The present sequence represents a TWEAK protein. The specification describes a method for preventing or treating an immunological disorder and/or inhibiting an immune response in an animal. The method comprises administering a TWEAK blocking agent. The method may be used for preventing and treating immune disorders associated with inappropriate expression and/or activity of TWEAK. These disorders include autoimmune diseases, acute and chronic inflammation, organ transplant rejection, Graft-versus-Host disease (GVHD), lymphoid cell malignancies, septic and other forms of shock, loss of immune



CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
CC promote wound healing or tissue grafting, by promoting vascularisation,  
CC also to induce apoptosis for treating cancer and eliminating autoreactive  
CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
CC TREPA peptides can also be used to target cytotoxic agents or for  
CC affinity isolation of the corresponding receptor, the nucleic acid for  
CC which can be used to transform tumour cells to render them more  
CC responsive to TREPA and to screen for TREPA mimics.  
CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
CC vascularisation), inflammation or a wide range of autoimmune conditions,  
CC conditions involving abnormal stimulation of epithelial cells (e.g.  
CC atherosclerosis), for birth control (inhibiting ovulation and placental  
CC formation) or other angiogenic conditions (e.g. ulcers).  
CC  
SQ Sequence 189 AA;  
Query Match 54.8%; Score 792; DB 19; Length 189;  
Best local Similarity 99.3%; Pred. No. 1.4e-67;  
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 132 KGRKTRARRAIAHYEHPRPGDGAQAGVDGTVSGMEARINSSPLRYNRQIGFIYV 191  
DB 37 KGRKTRARRAIAHYEHPRPGDGAQAGVDGTVSGMEARINSSPLRYNRQIGFIYV 96  
QY 192 RAGLYTYLCOVHDEGKAVYKLDLVGVIALRCLEEFSAATASSIGPOLRLCOVSGLL 251  
DB 97 RAGLYTYLCOVHDEGKAVYKLDLVGVIALRCLEEFSAATASSIGPOLRLCOVSGLL 156  
QY 252 ALRPGSSLRIRITLPMHAKAPFLTYFGLEQVH 284  
DB 157 ALRPGSSLRIRITLPMHAKAPFLTYFGLEQVH 189  
RESULT 14  
AAE00892  
ID AAE00892 standard; Protein; 189 AA.  
XX  
AC AAE00892;  
XX  
DT 04-JUL-2001 (first entry)  
XX  
DE Human UL4flag TREPA soluble construct.  
XX  
KM Human: tumour necrosis factor; TNF; angiogenesis; wound healing; tumour;  
KW TREPA; TNF related endothelium proliferative agent; metastasis; grafting;  
KM vulnery; HUVEC; human umbilical vein endothelial cell; UL4flag.  
XX  
OS Homo sapiens.  
XX  
PN US6207642-B1.  
XX  
PD 27-MAR-2001.  
XX  
PF 26-JUN-1998; 98US-0105343.  
XX  
PR 12-FEB-1997; 97US-0798692.  
PR 10-FEB-1998; 98US-0021706.  
XX  
PA (ABBO ) ABBOTT LAB.  
XX  
PI Wiley SR;  
XX  
DR WPI; 2001-280760/29.  
XX  
XX Inducing angiogenesis in mammal at desired sites for promoting wound  
PT healing, by administering soluble fragment of extracellular domain of  
PT tumor necrosis factor related endothelium proliferative agent protein  
PS Example 2; Column 75-78; 53pp; English.  
CC The present invention relates to extracellular signal molecules,

CC particularly members of tumour necrosis factor (TNF) family molecules  
CC designated as TREPA (TNF related endothelium proliferative agent).  
CC Soluble biologically active TREPA are used to treat TREPA-associated  
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
CC in human for promoting wound healing and for vascularising grafted  
CC tissue for successful grafting and to promote tissue grafts. The present  
CC amino acid sequence is human UL4flag TREPA soluble construct. This  
CC sequence which is a biologically active molecule is capable of inducing  
CC proliferation in HUVEC (human umbilical vein endothelial cells) cells.  
XX  
SQ Sequence 189 AA;  
Query Match 54.8%; Score 792; DB 22; Length 189;  
Best local Similarity 99.3%; Pred. No. 1.4e-67;  
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 132 KGRKTRARRAIAHYEHPRPGDGAQAGVDGTVSGMEARINSSPLRYNRQIGFIYV 191  
DB 37 KGRKTRARRAIAHYEHPRPGDGAQAGVDGTVSGMEARINSSPLRYNRQIGFIYV 96  
QY 192 RAGLYTYLCOVHDEGKAVYKLDLVGVIALRCLEEFSAATASSIGPOLRLCOVSGLL 251  
DB 97 RAGLYTYLCOVHDEGKAVYKLDLVGVIALRCLEEFSAATASSIGPOLRLCOVSGLL 156  
QY 252 ALRPGSSLRIRITLPMHAKAPFLTYFGLEQVH 284  
DB 157 ALRPGSSLRIRITLPMHAKAPFLTYFGLEQVH 189  
RESULT 15  
AAE00895  
ID AAE00895 standard; Protein; 146 AA.  
XX  
AC AAE00895;  
XX  
DT 04-JUL-2001 (first entry)  
XX  
DE Human TREPA (TNF related endothelium proliferative agent) fragment.  
XX  
KM Human: tumour necrosis factor; TNF; angiogenesis; wound healing;  
KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;  
KM grafting; vulnery.  
XX  
OS Homo sapiens.  
XX  
PN US6207642-B1.  
XX  
PD 27-MAR-2001.  
XX  
PF 26-JUN-1998; 98US-0105343.  
XX  
PR 12-FEB-1997; 97US-0798692.  
PR 10-FEB-1998; 98US-0021706.  
XX  
PA (ABBO ) ABBOTT LAB.  
XX  
PI Wiley SR;  
XX  
DR WPI; 2001-280760/29.  
XX  
XX Inducing angiogenesis in mammal at desired sites for promoting wound  
PT healing, by administering soluble fragment of extracellular domain of  
PT tumor necrosis factor related endothelium proliferative agent protein  
PS Example 14; Fig 1; 53pp; English.  
CC The present invention relates to extracellular signal molecules,  
CC particularly members of tumour necrosis factor (TNF) family molecules  
CC designated as TREPA (TNF related endothelium proliferative agent).  
CC Soluble biologically active TREPA are used to treat TREPA-associated  
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
CC in human for promoting wound healing and for vascularising grafted

CC tissue for successful grafting and to promote tissue grafts.  
 CC The present amino acid sequence is human TREPA fragment.  
 XX

Sequence 146 AA;

Query Match 52.7%; Score 761; DB 22; Length 146;  
 Best Local Similarity 100.0%; Pred. No. 9.3e-65;  
 Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 RRAIAHVEVHPRGDGAQAGVDGTVSGWEARINSSPLRYNRQIGEIFYTRAGLYYL 198  
 DB 1 RRAIAHVEVHPRGDGAQAGVDGTVSGWEARINSSPLRYNRQIGEIFYTRAGLYYL 60  
 QY 199 YCOVHDEGKAVYLYKIDLVGVLALRCLEFSATPAASSLGPOLRLCOVSGLALRPSS 258  
 DB 61 YCOVHDEGKAVYLYKIDLVGVLALRCLEFSATPAASSLGPOLRLCOVSGLALRPSS 120  
 QY 259 LRIRTLPMWHLKAPFLTYFGLFOVH 284  
 DB 121 LRIRTLPMWHLKAPFLTYFGLFOVH 146

Search completed: May 8, 2003, 02:19:42  
 Job time : 37.2672 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 05:32:15 ; Search time 58 Seconds  
(without alignments)  
1008.920 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 284  
Sequence: 1 MSLLDFEISARLPLPRSLG.....PMAHLKAPLTFYGLFQVH 284

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_protist:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	3.5	111	16	Q8X4J8
2	9	3.2	142	16	Q981J0
3	8	2.8	51	4	Q9BX52
4	8	2.8	143	17	Q9HST7
5	8	2.8	151	10	Q9SDI1
6	8	2.8	184	5	Q9VVT0
7	8	2.8	197	16	Q9WTU0
8	8	2.8	199	4	Q9BUI1
9	8	2.8	211	5	Q8SXH4
10	8	2.8	278	16	Q9ZG99
11	8	2.8	279	11	Q9D378
12	8	2.8	279	11	Q9CPR8
13	8	2.8	306	5	Q8SZB8
14	8	2.8	339	17	Q8S554
15	8	2.8	342	4	Q99908
16	8	2.8	343	4	Q9BPV2

17	8	2.8	370	17	Q8TRU8	Q8TUN8
18	8	2.8	372	16	Q53860	Q53860
19	8	2.8	377	10	Q9KXT3	Q9KXT3
20	8	2.8	387	10	Q9SHD8	Q9SHD8
21	8	2.8	397	16	Q8U820	Q8U820
22	8	2.8	435	16	Q8XS70	Q8XS70
23	8	2.8	443	4	Q9UFM5	Q9UFM5
24	8	2.8	454	5	Q9VNP0	Q9VNP0
25	8	2.8	465	16	Q8ZCV8	Q8ZCV8
26	8	2.8	471	16	Q9KR18	Q9KR18
27	8	2.8	472	4	Q96N66	Q96N66
28	8	2.8	473	11	Q9CY76	Q9CY76
29	8	2.8	473	11	Q8RIJ9	Q8RIJ9
30	8	2.8	522	10	Q9LW00	Q9LW00
31	8	2.8	564	12	Q9YTU9	Q9YTU9
32	8	2.8	568	12	Q8OPL0	Q8OPL0
33	8	2.8	656	5	Q9N8H2	Q9N8H2
34	8	2.8	1523	5	Q9N8U8	Q9N8U8
35	7	2.5	24	5	P83215	P83215
36	7	2.5	30	5	P83217	P83217
37	7	2.5	32	6	Q9BEF6	Q9BEF6
38	7	2.5	35	2	Q9ZG35	Q9ZG35
39	7	2.5	54	4	Q16193	Q16193
40	7	2.5	67	2	Q8ROS4	Q8ROS4
41	7	2.5	69	5	Q9VMK9	Q9VMK9
42	7	2.5	86	16	Q9AAG4	Q9AAG4
43	7	2.5	94	4	Q9P074	Q9P074
44	7	2.5	99	2	P97199	P97199
45	7	2.5	100	5	Q8WQR8	Q8WQR8

## ALIGNMENTS

## RESULT 1

Q8X4J8 PRELIMINARY; PRT; 111 AA.  
ID Q8X4J8  
AC Q8X4J8:  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Hypothetical protein z3516.  
GN z3516.  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OX NCBI\_TaxID=83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RX MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobbeck E.J., Davis N.W., Lam A., Dimalanta E.T., Potamousis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.:  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
RL Nature 409:529-533(2001).  
DR EMBL; AE005458; AAC57389.1; -  
DR InterPro; IPR000620; DUF6.  
DR Pfam; PF00893; DUF6; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 111 AA; 12165 MW; 7CEFC93D786CD759 CRC64;

Query Match 3.5%; Score 10; DB 16; Length 111;

Best Local Similarity 100.0%; Pred. No. 0.13; Mismatches 0; Indels 0; Gaps 0;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 LGLALACIGL 72  
Db 40 LGLALACIGL 49

RESULT 2  
 Q981J0 PRELIMINARY: PRT; 142 AA.  
 AC 0981J0: 01-OCT-2001 (TREMBlrel. 18, Created)  
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
 DE 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
 DE Probable transport protein.  
 GN MTR2380.  
 OS Rhizobium loti (Mesorhizobium loti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Phyllobacteriaceae; Mesorhizobium.  
 OX NCBI\_Taxid=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAFF303099;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
 Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,  
 Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
 Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabeta S.;  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 Mesorhizobium loti".  
 RT DML Res. 7:331-338(2000).  
 DR EMBL; AF002999; BAB49526.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 142 AA; 14864 MW; 0DCA7842CB5A56F CRC64;

Query Match 3.2%; Score 9; DB 16; Length 142;  
 Best Local Similarity 100.0%; Pred. No. 1.5;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 53 TALLVPLAL 61  
 |||||  
 DB 122 TALLVPLAL 130

RESULT 3  
 Q9BX52 PRELIMINARY: PRT; 51 AA.  
 AC 09BX52: 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE DJ831C21.1 (Novel protein similar to gamma-glutamyl  
 DE transpeptidase-related protein (GGF-Rel)) (Fragment).  
 GN DJ831C21.1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC Cory N.;  
 RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL133466; CAC34607.1; -.  
 FT NON\_TER 51 51  
 SQ SEQUENCE 51 AA; 5124 MW; 31980CCAE0451EF8 CRC64;

Query Match 2.8%; Score 8; DB 4; Length 51;  
 Best Local Similarity 100.0%; Pred. No. 5.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
 |||||  
 DB 7 LGIGLALA 14

RESULT 4  
 Q9HST7 PRELIMINARY: PRT; 143 AA.

AC Q9HST7: 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
 DE Vng0080H.  
 GN VNG0080H.  
 OS Halobacterium sp. (strain NRC-1).  
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
 OC Halobacteriaceae; Halobacterium.  
 OX NCBI\_Taxid=64091;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=20504483; PubMed=11016950;  
 RX Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,  
 RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Shroogna J.,  
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welt R., Goo Y.A.,  
 RA Leitner B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
 RA Isenbarger T.A., Beck R.F., Pohlischer M., Spudich J.L., Jung K.H.,  
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
 RT "Genome sequence of Halobacterium species NRC-1".  
 RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
 DR EMBL; AE004976; ANG18715.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 143 AA; 15648 MW; 45466E6328EF3468 CRC64;

Query Match 2.8%; Score 8; DB 17; Length 143;  
 Best Local Similarity 100.0%; Pred. No. 15;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68  
 |||||  
 DB 55 LGIGLALA 62

RESULT 5  
 Q9SD11 PRELIMINARY: PRT; 151 AA.  
 AC Q9SD11: 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Hypothetical protein (OSJNB0036E02.6 protein) (B1085F09.2  
 DE protein).  
 GN B1085F09.2.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_Taxid=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 RT clone:P0003H10.";  
 RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC  
 RT clone:OSJNB0036E02.";  
 RT Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC  
 RT clone:B1085F09.";  
 RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF000815; BAA87834.1; -.



**DR** EMBL: AP002862; BAB17732.1; -  
**OR** EMBL: AP003103; BAB4106.1; -  
**SO** SEQUENCE 151 AA; 16632 MW; EC68451ECA2BD71D CRC64;

	Query Match	Best Local Similarity	Score 8;	DB 10;	Length 151;
Matches	8;	Conservative	0;	Mismatches	0;
				Indels	Gaps
OY	43 RRRGRGE 50 				
Dd	131 RRRGRGE 138				

**RESULT 6**

ID	Q9VW70	PRELIMINARY;	PRT;	184 AA.	
AC	Q9VW70:				
DT	01-MAY-2000 (TREMblrel_13_Created)				
DT	01-MAY-2000 (TREMblrel_13_Last sequence update)				
DT	01-JUN-2002 (TREMblrel_21_Last annotation update)				
DE	CGI3033 protein.				
GN	CGI3033.				
OS	Drosophila melanogaster (Fruit fly).				
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Empidoidea; Drosophilidae; Drosophila.				
OX	NCBI_TaxID=7227;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-BERKELEY.				
FX	MEDLINE-20196006; PubMed-10731132. Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F., Gagne R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H., Blaise R.G., Champagne M., Pfeiffer B.D., Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abtill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P., Buttis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cavley S., Dahlke C., Davoport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M., Dodson K., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W., Foister C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K., Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-R., Idegawa C., Jaislaid M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Lin X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D., Mekulov G., Milshina N.V., Mobarry C., Morris J.J., Mosheriff A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Peterson K.A., Nixon K., Nusser D.R., Pacle J.M., Palazzo M., Platan G.S., Pan S., Pollard J., Puji V., Reese M.G., Rehnert K., Remington K., Saunders R.D.C., Schebler F., Shen H., Shue B.C., Sidlenklamos I., Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J., Williams S.M., Woodege T., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Yeh R.-F., Zavet J.S., Zhao M., Zhang G., Zhao O., Zheng L., Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; The genome sequence of Drosophila melanogaster." ; RL EMBL: AE003527; AAP94942.1; - OR FlyBase; FBgn0036638; CGI3033 OR InterPro; IPRO04019; YLP_motif. OR Pfam; PF02757; YLP; 5.				

```

SO SEQUENCE 184 AA; 21022 MW; 31976AE350DC447 CRC64;
Query Match 2.8%; Score 8; DB 5; Length 184;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 LLLAVSL 79
   |||||
DB 16 LLLAVSL 23

RESULT 7
Q9WYU0 PRELIMINARY; PRT; 197 AA.
ID Q9WYU0
AC Q9WYU0;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical protein TM0469.
GN TM0469.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Hat D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima.";
RL Nature 399:333-339(1999).
DR EMBL; AE001724; AAD35533.1; -.
DR TIGR; TM0469; -.
KW Hypothetical protein; Complete proteome.
SO SEQUENCE 197 AA; 22919 MW; 41B2C8E3C09180EC CRC64;

Query Match 2.8%; Score 8; DB 16; Length 197;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SLDPFIS 9
   |||||
DB 135 SLDPFIS 142

RESULT 8
Q9BUT1 PRELIMINARY; PRT; 199 AA.
ID Q9BUT1
AC Q9BUT1;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 22.8 kDa protein.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002598; AA02598.1; -.
DR InterPro; IPR004299; MBOAT_fam.
DR Pfam; PF03062; MBOAT; 1.
KW Hypothetical protein.
SO SEQUENCE 199 AA; 22750 MW; E5DAC747E7BEF06B CRC64;

```

Query Match 2.8%; Score 8; DB 4; Length 199;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLGIAL 67  
 DB 167 ALGLGIAL 174

RESULT 9  
 O8SXH4 PRELIMINARY; PRT; 211 AA.  
 AC O8SXH4;  
 DT 01-JUN-2002 (TEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)  
 DE RE50345P.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_Taxid=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RA Stapleton M., Brokstein P., Hong L., Agbayan A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,  
 RA Miranda A., Mungall C.J., Nuncio J., Pacleb J., Pargass V., Park S.,  
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
 RA Celisner S.,  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY089634; AAL90372.1;  
 SQ SEQUENCE 211 AA; 23780 MW; 82FP4983E91F510A CRC64;

Query Match 2.8%; Score 8; DB 5; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVSL 79  
 DB 9 LLLAVSL 16

RESULT 10  
 O9ZG99 PRELIMINARY; PRT; 278 AA.  
 AC O9ZG99;  
 DT 01-MAY-1999 (TEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)  
 DE Transmembrane protein AMPE.  
 GN AMPE OR PA4521.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 NCBI\_Taxid=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RA Langae Y.T., Dargis M., Huletsky A.;  
 RT "An amp gene in Pseudomonas aeruginosa encodes a negative regulator  
 of amp beta-lactamase expression."  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]

SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RC MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrone M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lartig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 opportunistic pathogen";  
 RL Nature 406:959-964(2000).  
 DR EMBL: AF082575; AAC98784.1; -;  
 DR EMBL: AE004866; AAG07909.1; -;  
 KW Transmembrane; Complete proteome.  
 SQ SEQUENCE 278 AA; 30793 MW; C623F1AB0691CFEF CRC64;

Query Match 2.8%; Score 8; DB 16; Length 278;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLAVY 77  
 DB 47 LGLLAVY 54

RESULT 11  
 O9D378 PRELIMINARY; PRT; 279 AA.  
 AC O9D378;  
 DT 01-JUN-2001 (TEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)  
 DE 5730494G16R1K protein.  
 GN 5730494G16R1K.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=MEDULLA OBLONGATA;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai T., Shingawa A., Shida K., Yoshino M., Itch M., Ishii Y.,  
 RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fletschmann W., Gaasterland T., Gissi C., King B., Kochava H.,  
 RA Kuehl P., Lewis S., Matsuo T., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Kombeerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection";  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK018250; BAB31133.1; -;  
 DR MGD: MGI:1913897; 5730494G16R1K.  
 DR InterPro: IPR002190; MAGE.  
 DR Pfam: PF01454; MAGE; 1.  
 SQ SEQUENCE 279 AA; 31474 MW; 5E243590A99F15F0 CRC64;

Query Match 2.8%; Score 8; DB 11; Length 279;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 80 GSRASLSA 87  
 DB 41 GSRASLSA 48

RESULT 12  
 O9CPR8 PRELIMINARY; PRT; 279 AA.  
 ID O9CPR8

Query	Db	80 GSRASLSA 87	41 GSRASLSA 48
RESULT 13			
08S2B8			
ID	08S2B8	PRELIMINARY;	PRT; 306 AA.
AC	08S2B8;		
DT	01-JUN-2002 (Tremblrel. 21, Created)		
DT	01-JUN-2002 (Tremblrel. 21, Last sequence update)		
DE	01-JUN-2002 (Tremblrel. 21, Last annotation update)		
DE	RE07882P.		
GN	CG1169.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_Taxid=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CG1BL/6J; TISSUE=EMBRYO, AND EMBRYONIC STEM CELLS.		
RC	MEDLINE=1085660; PubMed=1121785;		
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,		
RA	Atawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,		
RA	Sato T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batlow S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gliss C., King B., Kochia H.,		
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,		
RA	Schirral L.M., Staabli F., Suzuki K., Tomita M., Wagner L., Washio T.,		
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,		
RA	Blake J., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,		
RA	Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,		
RA	Guatlichin S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,		
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,		
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,		
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,		
RA	Suzuki H., Toyo-oka K., Wang K.H., Welter C., Whitaker C., Wilming L.,		
RA	Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawai J., Kohsaki S.,		
RA	Hayashizaki Y.;		
RT	"Functional annotation of a full-length mouse cDNA collection.";		
RL	Nature 409:685-690(2001).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RP	Aguirre P.H., Chomez P.M., De Backer O.R., Bertrand M.J.M.;		
RT	"Ten new murine members of the MAGE gene family."		
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AK017727; BAB30899.1; -		
DR	EMBL: AK010294; BAB26830.1; -		
DR	EMBL: AF319879; AKR01207.1; -		
DR	MCD; MG11913897; 5730494G16RLX.		
DR	InterPro: IPR002190; MAGE.		
DR	Pfam: PF01454; MAGE. 1.		
SO	SEQUENCE 279 AA; 31460 MW; FE2435919BD63160 CRC64;		
Query Match	2.8%; Score 8; DB 11; Length 279;		
Best Local Similarity	100.0%; Pred. No. 26;		
Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		

RA	Stapleton M., Brokstein P., Hong L., Abdayani A., Carlson J.,
RA	Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA	George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA	Miranda A., Munzall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA	Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA	Celniker S.; Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
DR	EMBL; AY070962; AL48604.1; -
SQ	SEQUENCE 306 AA; 34083 MW; 32B69371475A48F9 CRC64;
OY	Best local Similarity 2.8%; Score 8; DB 29; Length 306;
Db	Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
	170 LVPLALGL 177
RESULT 14	
ID	058554 PRELIMINARY; PRT; 339 AA.
AC	058554;
DT	01-AUG-1998 (TREMBlrel. 07, Created)
DT	01-JAN-1999 (TREMBlrel. 09, Last sequence update)
DT	01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE	Hypothetical protein PH0824.
GN	PH0824.
OS	Pyrococcus horikoshii.
OC	Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC	Pyrococcus.
OX	NCBI_Taxid=53953;
RN	[1]
RP	SEQUENCE FROM N.A..
RC	STRAIN=OT3;
RX	MEDLINE=98344137; PubMed=9679194;
RA	Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA	Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA	Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA	Funahshi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,
RA	Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA	Masuchi Y., Shitaya H., Kikuchi H.;
RT	"Complete sequence and gene organization of the genome of a hyper-
RT	thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL	DNA Res. 5:55-76(1998).
DR	EMBL; AP000003; BAA29917.1; -
DR	InterPro: IPR002657; BiLeac/Na_smprtr.
KW	Pfam: PF01758; SBF; 1.
KW	Hypothetical protein; Complete proteome.
SQ	SEQUENCE 339 AA; 37228 MW; E91697D5C8C3705F CRC64;
OY	Query Match 2.8%; Score 8; DB 17; Length 339;
Db	Best Local Similarity 100.0%; Pred. No. 31;
	Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
	70 LGLLAVV 77
RESULT 15	
ID	099908 PRELIMINARY; PRT; 342 AA.
AC	099908;
DT	01-MAY-1997 (TREMBlrel. 03, Created)
DT	01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT	01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE	BH1 protein.
GN	BH1.
OS	Homo sapiens (Human).
OC	Eukaryota; Metacoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_Taxid=9606;	

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-96273128; PubMed-8702217;
RA Fukunaga-Johnson N., Lee S.W., Liebert M., Grossman H.B.;
RT "Molecular analysis of a gene, BBI, overexpressed in bladder and
  breast carcinoma.";
RL Anticancer Res. 16:1085-1090(1996).
DR EMBL: S82470; AAB37433.1; -
DR InterPro: IPR004299; MBOAT_fam.
SQ SEQUENCE 342 AA; 38163 MW; 2B479EABCF1B91C CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 32; Length 342;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 ALGGLAL 67
Db 310 ALGGLAL 317

RESULT 16
Q9BPV2 PRELIMINARY; PRT; 343 AA.
ID Q9BPV2
AC Q9BPV2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Unknown (protein for MGC:4221) (protein for MGC:2099).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=KIDNEY, AND COLON;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC003164; AA003164.1; -
DR EMBL: BC002512; AA002512.1; -
DR InterPro: IPR004299; MBOAT_fam.
DR Pfam: PF03062; MBOAT; 1.
SQ SEQUENCE 343 AA; 38727 MW; F71E7DBF74BD9B7 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 32; Length 343;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 ALGGLAL 67
Db 311 ALGGLAL 318

RESULT 17
Q8TUB8 PRELIMINARY; PRT; 370 AA.
ID Q8TUB8
AC Q8TUB8;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Permease subunit of a ABC-type transport system involved in
  lipid protein release.
GN MK1655.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.
OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-AV19 / DSM 6324 / JCM 9639;
RX MEDLINE-21927647; PubMed-11930014;
RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
  Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,

```

```

RA Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
  Malikh A.G., Koonin E.V., Kozzyavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
  and monophyly of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR EMBL: AE010455; AA02868.1; -
RW Complete proteome.
SQ SEQUENCE 370 AA; 39411 MW; B07662EA1E5A644E CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 34; Length 370;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LGGLALA 68
Db 336 LGGLALA 343

RESULT 18
O53860 PRELIMINARY; PRT; 372 AA.
ID O53860
AC O53860;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein cysM3.
GN CYSM3 OR RV0848 OR MT043. 41.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE-98295987; PubMed-9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
  Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
  Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
  Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
  Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
  Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
  Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
  Sutcliffe J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
  complete genome sequence.";
RL Nature 393:537-544(1998).
DR EMBL: AL022004; CAAL17654.1; -
DR HSP: P35520; 1JBO.
DR Tuberculist; RV0848; -
DR InterPro: IPR001926; B6_enzyme_beta.
DR Pfam: PF00291; PALP; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 372 AA; 40118 MW; 927386B1DF5F8C CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 34; Length 372;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LGGLALA 68
Db 96 LGGLALA 103

RESULT 19
O9KXT3 PRELIMINARY; PRT; 377 AA.
ID O9KXT3
AC O9KXT3;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative integral membrane protein.
GN SC05682 OR SC5H4.06C.
OS Streptomyces coelicolor.

```

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.  
 OX NCBI\_Taxid=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2) / M45;  
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Godle A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Metzorroek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RT coelicolor A3(2).";  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL355913; CAB91118.1;  
 SQ SEQUENCE 377 AA; 37614 MW; A35DA0437F04AC46 CRC64;  
 Query Match 2.8%; Score 8; DB 16; Length 377;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 59 LALGGLA 66  
 |||||  
 Db 327 LALGGLA 334  
 RESULT 20  
 O9SHD8 PRELIMINARY; PRT: 387 AA.  
 ID O9SHD8;  
 AC O9SHD8;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created).  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE At2g45000 protein (At2g45000/T14P1.20) (Hypothetical 40.6 kDa  
 DE protein).  
 GN AT2G45000 OR AT2G45005.  
 GN Arabidopsis thaliana (Mouse-ear cress).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RX MEDLINE=20083487; PubMed=10617197;  
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Bentto M.-I., Town C.D.,  
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,  
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,  
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,  
 RA Coppenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,  
 RA Salberg S.L., Fraser C.M., Venter J.C.;  
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis  
 RT thaliana.";  
 RL Nature 402:761-768(1999).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CV. COLUMBIA;  
 RA Lin X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Shin P., Chen H., Cheuk R., Kim C.J., Lim J., Meyers M.C., Banh J.,  
 RA Bower L., Carinci P., Chang E., Dale J.M., Goldsmith A.D.,  
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Kalin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Miranda H.L., Sakurai T., Satou M.,  
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,

RA Ecker J.R.;  
 RT "Arabidopsis cDNA clones.";  
 RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
 RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bower L.,  
 RA Carinci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
 RA Kamiya A., Kalin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,  
 RA Seki M., Shin P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
 RA Theologis A.;  
 RT "Arabidopsis Full Length cDNA Clones.";  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC007659; AAD32835.1;  
 DR EMBL: AY074646; AAL69462.1;  
 DR EMBL: AY080620; AAL6303.1;  
 KW Hypothetical protein.  
 SQ SEQUENCE 387 AA; 40584 MW; AF6C6B3BAC9BP69A CRC64;  
 Query Match 2.8%; Score 8; DB 10; Length 387;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 93 EELVAED 100  
 |||||  
 Db 370 EELVAED 377  
 RESULT 21  
 O8U820 PRELIMINARY; PRT: 397 AA.  
 ID O8U820;  
 AC O8U820;  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hypothetical protein Atu3948.  
 GN AtU3948 OR AGR.L.1808.  
 GN Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OS Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Rhizobium.  
 OX NCBI\_Taxid=176299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608550; PubMed=11743193;  
 RA Wood D.W., Setubal J.C., Kaul R., Marks D.E., Kitajima J.P.,  
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,  
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
 RA Kutayavin T., Levy R., Li M.-J., McLelland E., Palmeri A.,  
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.;  
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens  
 RT C58.";  
 RL Science 294:2317-2323(2001).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608551; PubMed=11743194;  
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Gurolo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Hounlet K., Gordon J., Vaudin M., Tatchouk O., Epp A., Liu F.,  
 RA Wollam C., Allinger M., Doughy D., Scott C., Lappas C., Martelz B.,  
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Seear C., Strub G.,  
 RA Gioe C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 RT Agrobacterium tumefaciens C58.";  
 RL Science 294:2323-2328(2001).  
 DR EMBL: AE009325; AAL44750.1; ALT\_INIT.  
 DR EMBL: AE008289; AAK89478.1;  
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 397 AA; 41708 MW; 700748E32A46AE86 CRC64;  
 Query Match 2.8%; Score 8; DB 16; Length 397;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 251 LALRPGSS 258  
 |||||  
 Db 341 LALRPGSS 348

RESULT 22  
 Q8XS70 PRELIMINARY; PRT; 435 AA.  
 AC Q8XS70;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 GN Probable transmembrane protein.  
 OS RSP0611 OR RS03756.  
 ON Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Plasmid megaplasmid.  
 CC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
 OC Ralstonia.  
 OX NCBI\_TaxID=305;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GM11000;  
 RX MEDLINE-21681879; PubMed-11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,  
 RA Arlet M., Billault A., Brotlier P., Camus J.C., Catolico L.,  
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Sautin W., Schlex T.,  
 RA Sigler P., Trebault P., Whalen M., Wincker P., Levy M.,  
 RA Weissenbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum";  
 RL Nature 415:497-502(2002).  
 DR EMBL: AL646079; CAD17762.1; -  
 DR InterPro: IPR001564; NDK.  
 DR PROSITE: PS00469; NDP\_KINASES; UNKNOWN\_1.  
 KM Plasmid: Complete proteome.  
 SQ SEQUENCE 435 AA; 47048 MW; CCD859D9C54DDB5A CRC64;

Query Match 2.8%; Score 8; DB 16; Length 435;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 217 LVDGVIAL 224  
 |||||  
 Db 358 LVDGVIAL 365

RESULT 23  
 Q9UFM5 PRELIMINARY; PRT; 443 AA.  
 AC Q9UFM5;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 GN Hypothetical 47.4 kDa protein..  
 OS DREFP5660011.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-KIDNEY;  
 RA Blum H., Bauersachs S., Mewes H.W., Gaassenhuber J., Wiemann S.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL117143; CAB55910.1; -  
 DR MEROPS: T03.002; -  
 DR InterPro: IPR00101; Gglutnspptase.

DR Pfam: PF01019; G\_glu\_transpept; 1.  
 DR PRINTS: PR01210; GGTTRANSPTASE.  
 KW Hypothetical protein.  
 SQ SEQUENCE 443 AA; 47446 MW; 3D75E9DB08265971 CRC64;

Query Match 2.8%; Score 8; DB 4; Length 443;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGGLALA 68  
 |||||  
 Db 14 LGGLALA 21

RESULT 24  
 Q9VNP0 PRELIMINARY; PRT; 454 AA.  
 AC Q9VNP0;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 GN CG1169 protein.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferrar C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Klumel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liao X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,  
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Palenart K., Remington K., Saunders R.D.C., Scheeler F., Shen R.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weisstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003600; AAF51889.1; -  
 DR FLYBASE: FBgn0037428; CG1169.  
 SQ SEQUENCE 454 AA; 51320 MW; A75AAAD97E716573 CRC64;

Query Match  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 LVPALGL 63  
170 LVPALGL 177

RESULT 25  
08ZCV8 PRELIMINARY; PRT; 465 AA.  
ID 08ZCV8  
AC 08ZCV8;  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Putative membrane protein.  
GN YECG OR YPO2850.  
OS Yersinia pestis.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Yersinia  
OX NCBI\_TaxID=632;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CO-92 / BIOVAR ORIENTALIS;  
RX MEDLINE=21470413; PubMed=11586360;  
RA Pakhilli J., Wren B.W., Thomson N.R., Tiltball R.W., Holden M.T.G.,  
RA Prentice M.B., Sebaldia M., James K.D., Churcher C., Mungall K.L.,  
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,  
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,  
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.,  
RT "Genome sequence of Yersinia pestis, the causative agent of plague."  
RL Nature 413:523-527(2001).  
DR EMBL: AJ14154; CAC92102.1; -  
DR InterPro: IPR003662; sub transporter.  
DR Pfam: PF00083; sugar\_tr.1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 465 AA; 50176 MW; OCC273F10B83F5ED CRC64;

Query Match  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 72 LLLAVSL 79  
335 LLLAVSL 342

RESULT 26  
09RR18 PRELIMINARY; PRT; 471 AA.  
ID 09RR18  
AC 09RR18;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE Transport protein, putative.  
GN DR2502.  
OS Deinococcus radiodurans.  
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;  
OC Deinococcaceae; Deinococcus.  
OX NCBI\_TaxID=1299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=20036896; PubMed=10567266;  
RX White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
RA Motil K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,  
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,

RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
RA Fraser C.M.;  
RT "Genome sequence of the radioresistant bacterium Deinococcus  
RT radiodurans R1."  
RL Science 286:1571-1577(1999).  
DR EMBL: AE002079; AAF12043.1; -  
DR TIGR: DR2502; -  
KW Complete proteome.  
SQ SEQUENCE 471 AA; 47974 MW; 96B2B8BFEE445D27 CRC64;

Query Match  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 ALGLGL 67  
366 ALGLGL 373

RESULT 27  
096N66 PRELIMINARY; PRT; 472 AA.  
ID 096N66  
AC 096N66;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE CDNA FLJ131346 f1s, clone MESAN1000180, highly similar to BBI-malignant  
DE cell expression-enhanced gene/tumor progression-enhanced gene.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Taahiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,  
RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,  
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,  
RA Kawachi-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,  
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,  
RA Murakawa K., Kanemori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,  
RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K.,  
RA Iisaga T.;  
RT "NEDO human cDNA sequencing project."  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AK055908; BAB71043.1; -  
DR InterPro: IPR004299; MEOAT\_fam.  
DR Pfam: PF03062; MEOAT; 1.  
SQ SEQUENCE 472 AA; 52774 MW; EA721998043F9EBD CRC64;

Query Match  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 ALGLGL 67  
440 ALGLGL 447

RESULT 28  
09CY76 PRELIMINARY; PRT; 473 AA.  
ID 09CY76  
AC 09CY76;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE 5730589L02RIK protein.  
GN 5730589L02RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN-C57BL/6J; TISSUE-EMBRYO;  
 RX MEDLINE-21085660; PubMed-11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,  
 RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleschmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,  
 RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RT Nature 409:685-690(2001).  
 DR EMBL: AK019981; BAB31950.1; -;  
 DR MGI:1924832; 5730589L02Rik.  
 DR InterPro: IPR004239; MBOAT\_Fam.  
 DR Pfam: PF03062; MBOAT; 1.  
 SO SEQUENCE 473 AA; 53504 MW; CE6FE893C3D01C4F CRC64;

Query Match 2.8%; Score 8; DB 11; Length 473;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACGL 72  
 Db 436 LALACGL 443

RESULT 29  
 O8R1P9 PRELIMINARY; PRT; 473 AA.  
 AC O8R1P9;  
 DT 01-JUN-2002 (Tremblrel. 21, Created)  
 DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
 DE RIKEN cDNA 5730589L02 gene.  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC023417; AAH23417.1; -;  
 SO SEQUENCE 473 AA; 53382 MW; DAA1FEDDA78013EA CRC64;

Query Match 2.8%; Score 8; DB 11; Length 473;  
 Best Local Similarity 100.0%; Pred. No. 42;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACGL 72  
 Db 436 LALACGL 443

RESULT 30  
 O9L1W0 PRELIMINARY; PRT; 522 AA.  
 AC O9L1W0;  
 DT 01-OCT-2000 (Tremblrel. 15, Created)  
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
 DT 01-OCT-2000 (Tremblrel. 15, Last annotation update)  
 DE Similar to an Arabidopsis thaliana chromosome BAC genomic

DE sequence.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hsing Y.C., Chow T., Chen C., Wu H., Chu M., Chao Y., Liu S.;  
 RT "Oryza sativa PAC P0699E04 genomics sequence, complete sequence.";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AP01111; BAA90509.1; -;  
 SO SEQUENCE 522 AA; 54697 MW; 21C6BAD2441B56BF CRC64;

Query Match 2.8%; Score 8; DB 10; Length 522;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRGE 50  
 Db 415 RRRGRGE 422

RESULT 31  
 O9YTU9 PRELIMINARY; PRT; 564 AA.  
 AC O9YTU9;  
 DT 01-MAY-1999 (Tremblrel. 10, Created)  
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
 DE Hemagglutinin H5 (fragment).  
 OS Influenzavirus A.  
 OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
 OC Influenza A viruses.  
 OX NCBI\_TaxID=11320;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A/DUCK/PODSAM/2216-4/84;  
 RX MEDLINE-99099002; PubMed-9882316;  
 RA Matrosovich M., Zhou N., Kawaoke Y., Webster R.;  
 RT "The surface glycoproteins of H5 influenza viruses isolated from humans, chickens, and wild aquatic birds have distinguishable properties.";  
 RT J. Virol. 73:1146-1155(1999).  
 CC - FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO  
 CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).  
 CC - SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS  
 CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).  
 CC - SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.  
 DR EMBL: AF082041; AAD13573.1; -;  
 DR HSSP: P03437; IHTM.  
 DR InterPro: IPR001364; Hemagglutn.  
 DR Pfam: PF00509; Hemagglutinin; 1.  
 DR PRINTS: PRO0329; HEMAGGLUTIN12.  
 DR ProDom: PD000225; Hemagglutn; 1.  
 KW Envelope protein; Glycoprotein; Hemagglutinin.  
 FT NON\_TER 564  
 FT 564  
 SO SEQUENCE 564 AA; 63562 MW; B317179A7E3E6F98 CRC64;

Query Match 2.8%; Score 8; DB 12; Length 564;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVSL 79  
 Db 6 LLLAVSL 13

RESULT 32  
 O8QPL0 PRELIMINARY; PRT; 568 AA.  
 AC O8QPL0;  
 DT 01-JUN-2002 (Tremblrel. 21, Created)



DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)  
 DE 01-JUN-2002 (TREMBLrel. 21, last annotation update)  
 GN Hemagglutinin (Fragment).  
 OS Influenza A virus (A/Goose/Hong Kong/3014.8/2000(H5N1)).  
 OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
 OC Influenza A viruses; Influenzavirus A.  
 OX NCBI\_TaxID=176675;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A/GOOSE/HONG KONG/3014.8/2000(H5N1);  
 RA MEDLINE=21874832; PubMed=11878904;  
 RA Guan Y., Peiris M., Kong K.F., Dyrting K.C., Ellis T.M., Sit T.,  
 RA Zhang L.J., Shortridge K.F.;  
 RT \*H5N1 Influenza Viruses Isolated from Geese in Southeastern China:  
 RT Evidence for Genetic Reassortment and Interspecies Transmission to  
 RT Ducks.\*;  
 RL Virology 292:16-23(2002).  
 DR EMBL; AY059482; AAL31388.1; .  
 FT NON\_TER 568 568  
 SQ SEQUENCE 568 AA; 64281 MW; 0B0A4CE034F1769 CRC64;  
  
 Query Match 2.8%; Score 8; DB 12; Length 568;  
 Best Local Similarity 100.0%; Pred. No. 50;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 72 LLLAVSL 79  
 DB 6 LLLAVSL 13  
  
 RESULT 33  
 ID Q9N8H2 PRELIMINARY; PRT; 656 AA.  
 AC Q9N8H2;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
 DE 01-OCT-2000 (TREMBLrel. 15, last annotation update)  
 DE Possible hypothetical 61.9 kDa protein.  
 GN CHRI.338.  
 OS Trypanosoma brucei.  
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.  
 OX NCBI\_TaxID=5691;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FRE927;  
 RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,  
 RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,  
 RA Gerard C., Rajandream M.A., Barrell B.G.;  
 RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AL359782; CAB95571.1; .  
 FT NON\_TER 656 656  
 SQ SEQUENCE 656 AA; 72138 MW; CBAC892D25937FAD CRC64;  
  
 Query Match 2.8%; Score 8; DB 5; Length 656;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 93 EELVAED 100  
 DB 454 EELVAED 461  
  
 RESULT 34  
 ID Q9N8U8 PRELIMINARY; PRT; 1523 AA.  
 AC Q9N8U8;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
 DE Possibly hypothetical protein 85.6 kD.  
 GN CHRI.139.  
 OS Trypanosoma brucei.  
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.

OX NCBI\_TaxID=5691;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FRE927;  
 RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,  
 RA Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,  
 RA Gerard C., Rajandream M.A., Barrell B.G.;  
 RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AL359782; CAB95435.1; .  
 FT NON\_TER 1523 1523  
 SQ SEQUENCE 1523 AA; 168322 MW; 07BDC751CDD1E5A CRC64;  
  
 Query Match 2.8%; Score 8; DB 5; Length 1523;  
 Best Local Similarity 100.0%; Pred. No. 1,2e+02;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 218 VDGVALR 225  
 DB 382 VDGVALR 389  
  
 RESULT 35  
 ID P83215 PRELIMINARY; PRT; 24 AA.  
 AC P83215;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
 DE Sperm protamine p3 (Po3) (Fragment).  
 OS Octopus vulgaris (Octopus).  
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;  
 OC Incurrata; Octopodidae; Octopus.  
 OX NCBI\_TaxID=6645;  
 RN [1]  
 RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
 RC TISSUE=SPERM;  
 RA Gimenez-Bonafe P., Ribes E., Buesa C., Sautiere P., Kouach M.,  
 RA Ausio J., Kasinsky H.E., Chiva M.;  
 RT \*Chromatin Remodelling and Protamines during Spermiogenesis of Octopus  
 RT vulgaris (Cephalopoda).\*;  
 RL J. Exp. Zool. 0:0-0(2001).  
 CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF  
 CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT  
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- TISSUE SPECIFICITY: TESTIS.  
 CC -1- MASS SPECTROMETRY: MW=4389; METHOD=ELECTROSPRAY.  
 KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;  
 FT DNA condensation; Nuclear protein.  
 FT DOMAIN 1 16 POLY-ARG.  
 FT NON\_TER 24 24  
 SQ SEQUENCE 24 AA; 3381 MW; 308E90ED9D2C9C9C CRC64;  
  
 Query Match 2.5%; Score 7; DB 5; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
 QY 43 RRRGRG 49  
 DB 11 RRRGRG 17  
  
 RESULT 36  
 ID P83217 PRELIMINARY; PRT; 30 AA.  
 AC P83217;  
 DT 01-MAR-2002 (TREMBLrel. 20, Created)  
 DT 01-MAR-2002 (TREMBLrel. 20, last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)  
 DE Sperm protamine p5 (Po5).  
 OS Octopus vulgaris (Octopus).  
 OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;  
 OC Incurrata; Octopodidae; Octopus.

OX NCBI\_TaxID=6645;  
 RN [1]  
 RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
 RC TISSUE-SPERM;  
 RA Gimenez-Bonafé P., Ribes E., Buesa C., Sautiere P., Kouach M.,  
 RA Ausio J., Kasinsky H.E., Chliva M.;  
 RT "Chromatin remodeling and protamines during spermatogenesis of Octopus  
 RT vulgaris (Cephalopoda).";  
 RL J. Exp. Zool. 0:0-0(2001).  
 CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF  
 CC SPERM DURING THE HAPOID PHASE OF SPERMATOGENESIS. THEY COMPACT  
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- TISSUE SPECIFICITY: TESTIS.  
 CC -1- MASS SPECTROMETRY: MW=3941; METHOD=ELECTROSPRAY.  
 KM Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;  
 FT DOMAIN 2 15 POLY-ARG.  
 FT DOMAIN 17 26 POLY-ARG.  
 SQ SEQUENCE 30 AA; 3943 MW; 14F1BC7E4D277049 CRC64;

Query Match 2.5%; Score 7; DB 5; Length 30;  
 Best Local Similarity 100.0%; Pred. No. 35;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRGRRG 49  
 DB 5 RRGRRG 11

RESULT 37  
 O9BEF6 PRELIMINARY; PRT; 32 AA.

DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Beta-lactoglobulin precursor (Fragment).  
 GN BETA-LACTOGLOBULIN.  
 OS Capra hircus (Goat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Capra.  
 OX NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA D'Andrea M., Pilla F., Graziano M.;  
 RT "A new polymorphism of goat beta-lactoglobulin proximal promoter  
 RT region.";  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ292058; CAC27455.1;  
 KW Signal.  
 FT SIGNAL 1 18 POTENTIAL.  
 FT NON\_TER 32 32  
 SQ SEQUENCE 32 AA; 3372 MW; 0C56BD579B3DC190 CRC64;

Query Match 2.5%; Score 7; DB 6; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGLALAC 69  
 DB 8 LGLALAC 14

RESULT 38  
 O9ZG35 PRELIMINARY; PRT; 35 AA.  
 ID O9ZG35  
 AC O9ZG35  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)  
 DE Hypothetical 3.5 kDa protein (Fragment).

OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-L2 434B;  
 RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;  
 RT "Gene identification of Chlamydia trachomatis by random DNA  
 RT sequencing.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF087333; AAD4107.1;  
 KW Hypothetical protein.  
 FT NON\_TER 1 1  
 FT NON\_TER 35 35  
 SQ SEQUENCE 35 AA; 3519 MW; 4686B72BAF28F7D CRC64;

Query Match 2.5%; Score 7; DB 2; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 40;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 NSSSPLR 180  
 DB 8 NSSSPLR 14

RESULT 39  
 O16193 PRELIMINARY; PRT; 54 AA.

DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE Orf2 5' to PD-ECGF/TP protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=94312438; PubMed=8038210;  
 RA USUKI K., GONZ L.J., WERNSTEDT C., MOREN A., MIYAZONO K.,  
 RA Claesson-Welsh L., Heldin C.H.;  
 RT "Structural properties of 3.0 kb and 3.2 kb transcripts encoding  
 RT platelet-derived endothelial cell growth factor/thymidine  
 RT phosphorylase in A431 cells.";  
 RL Biochim. Biophys. Acta 1222:411-414(1994).  
 DR EMBL: S72487; AAD14107.1;  
 SQ SEQUENCE 54 AA; 5808 MW; 827925FBA70CB222 CRC64;

Query Match 2.5%; Score 7; DB 4; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 125 RPRSSAP 131  
 DB 17 RPRSSAP 23

RESULT 40  
 O8ROS4 PRELIMINARY; PRT; 67 AA.  
 ID O8ROS4  
 AC O8ROS4  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hypothetical 7.8 kDa protein.  
 GN YHCR.  
 OS Serratia marcescens.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Serratia.  
 OX NCBI\_TaxID=615;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA Baba M., Midorikawa Y., Nakagawa Y., Fujita M., Matsuyama T.;  
 RT "Serratia marcescens and Escherichia coli genes controlling  
 RT temperature-dependent production of structurally unrelated secondary  
 RT metabolites such as prodigiosin and serrawettin.";  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB080601; BAB85653.1; -  
 KM Hypothetical protein.  
 SQ SEQUENCE 67 AA; 7791 MW; 47B01A87E69AC2A2 CRC64;

Query Match 2.5%; Score 7; DB 2; Length 67;  
 Best Local Similarity 100.0%; Pred. No. 72;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 LVLVPL 61  
 DB 21 LVLVPL 27

## RESULT 41

O9VMM9 PRELIMINARY; PRT; 69 AA.

AC 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE CG11147 protein.  
 GN CG11147.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 CC NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;  
 RX MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouch J., Brokstein P., Brotlier P.,  
 RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Duthin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,  
 RA Relier K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Glids R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";

RL Science 287:2185-2195 (2000).  
 DR EMBL: AE003611; AAF52284.1; -  
 DR Flybase: FBgn0031734; CG11147.  
 SQ SEQUENCE 69 AA; 8118 MW; 9728A36DDE94915F CRC64;

Query Match 2.5%; Score 7; DB 5; Length 69;  
 Best Local Similarity 100.0%; Pred. No. 74;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 42 QRRGR 48  
 DB 54 QRRGR 60

## RESULT 42

O9A4G4 PRELIMINARY; PRT; 86 AA.

AC 09A4G4;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)  
 DE Hypothetical protein CC2870.  
 GN CC2870.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 CC NCBI\_TaxID=155892;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 19089 / CB15;  
 RX MEDLINE-21173698; PubMed-11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Onda N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Debey R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Utterback T., Tran K., Wolf A., Vamathevan J., Emolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005952; AAK24834.1; -  
 DT TIGR: CC2870;  
 KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 86 AA; 7615 MW; EEF6383D9A545E CRC64;

Query Match 2.5%; Score 7; DB 16; Length 86;  
 Best Local Similarity 100.0%; Pred. No. 89;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 LALGGL 65  
 DB 26 LALGGL 32

## RESULT 43

O9P074 PRELIMINARY; PRT; 94 AA.

AC 09P074;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE HSPC308 (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 CC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE-BLOOD;  
 RA Ye M., Zhang Q.H., Zhou J., Shen Y., Wu X.Y., Guan Z.O., Wang L.,  
 RA Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;  
 RT "Human partial CDS from cd34+ stem cells.";  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.

```

DR EMBL: AF161426; AAF28986.1; -
FT NON_TER 94
SQ SEQUENCE 94 AA; 10874 MW; CDA7FD97C850842P CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 4; Length 94;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 EEDDPS 104
    |||||
Db 11 EEDDPS 17

RESULT 44
P97199 PRELIMINARY; PRT; 99 AA.
ID P97199;
AC P97199;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE Similar to.
GN IS5.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97251358; Pubmed=9097040;
RA Itoh T., Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Kasai H., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T.,
RA Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y.,
RA Nishimoto H., Nishio Y., Oshima T., Saito N., Sempel G., Seki Y.,
RA Sivasubraman S., Tagami H., Takeda J., Takemoto K., Wada C.,
RA Yamamoto Y., Horinouchi T.;
RT "A 460-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 40,150,000 min region on the linkage map.";
RL DNA Res. 3:379-392(1996).
DR EMBL: D90837; BAA15812.1; -
DR EMBL: D90838; BAA15817.1; -
SQ SEQUENCE 99 AA; 10142 MW; 03EBF8DD10E92C0 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 2; Length 99;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 DGAVRQ 29
    |||||
Db 88 DGAVRQ 94

RESULT 45
Q8MOR8 PRELIMINARY; PRT; 100 AA.
ID Q8MOR8;
AC Q8MOR8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Lox4 homeodomain protein (Fragment).
GN LOX4.
OS Euprymna scolopes.
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Sepioida;
OC Sepiidae; Euprymna.
OX NCBI_TaxID=6613;
RN [1]
RP SEQUENCE FROM N.A.
RA Calleerts P., Lee P.N., Hartmann B., Farfan C.B., Choy D.W.Y.,
RA Fischbach K.F., Gehring W.J., de Couet H.G.;
RT "HOX genes in the sepioid squid Euprymna scolopes: implications for
RT the evolution of complex body plans.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AT052759; AAL25810.1; -

```

```

DR InterPro; IPR001827; Antennapedia.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR000047; HTH_repressor.
DR Pfam; PF00046; homeobox; 1.
DR PRINTS; PR00025; ANTENNAPEDIA.
DR PRINTS; PR00024; HOMEBOX.
DR PRINTS; PR00031; HTHREPRESSR.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; UNKNOWN_1.
DR PROSITE; PS50071; HOMEBOX_2; 1.
KW Homeobox; DNA-binding; Nuclear protein.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 100 AA; 12053 MW; 26C262722E4DC514 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 5; Length 100;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 41 SQRGR 47
    |||||
Db 18 SQRGR 24

Search completed: May 8, 2003, 06:57:36
Job time : 62 secs

```

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 04:39:54 ; Search time 35 seconds

(without alignments)  
336.551 Million cell updates/sec

Title: US-09-245-198A-4

Perfect score: 284  
Sequence: 1 MSLLDFEISARRLPLPRSLG.....PMAHLKAPFLTYGELFQVH 284

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size: 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt.40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	249	87.7	249	1	TN12_HUMAN
2	32	11.3	225	1	TN12_MOUSE
3	10	3.5	111	1	YFWM_ECOLI
4	9	3.2	733	1	PSAB_ODOSI
5	8	3.2	734	1	PSAB_CVACA
6	8	2.8	179	1	ADHS_GLIOX
7	8	2.8	220	1	Y304_BROME
8	8	2.8	317	1	MSHR_PANTR
9	8	2.8	379	1	FDHB_METJA
10	8	2.8	422	1	ZP3_MESAU
11	8	2.8	576	1	CYDC_HAETN
12	8	2.8	586	1	GCT5_HUMAN
13	7	2.5	49	1	HSP1_SAGIM
14	7	2.5	57	1	HSP1_DIDMA
15	7	2.5	115	1	A62F_DROME
16	7	2.5	118	1	Y151_ECOLI
17	7	2.5	131	1	IL13_MOUSE
18	7	2.5	131	1	IL13_RAT
19	7	2.5	147	1	YK01_PANRO
20	7	2.5	150	1	TNFC_PIG
21	7	2.5	157	1	RA05_ORYSA
22	7	2.5	161	1	HLPA_ECOLI
23	7	2.5	170	1	Y1M4_CAEEL
24	7	2.5	180	1	LACB_BUBBU
25	7	2.5	180	1	LACB_CAPII
26	7	2.5	180	1	LACB_SHEEP
27	7	2.5	230	1	RNFE_SALTY
28	7	2.5	231	1	RNFE_ECOS7
29	7	2.5	231	1	RNFE_ECOLI
30	7	2.5	233	1	RNFE_ECOLI
31	7	2.5	233	1	RNFE_YERPE
32	7	2.5	235	1	RNFE_HAETN
33	7	2.5	239	1	TN14_MOUSE

## ALIGNMENTS

RESULT 1	ID	TN12_HUMAN	STANDARD:	PRT: 249 AA.
AC	043508	Q8WU27		
DT	15-JUN-2002	(Rel. 41, Created)		
DT	15-JUN-2002	(Rel. 41, Last sequence update)		
DT	15-JUN-2002	(Rel. 41, Last annotation update)		
DE	Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (APO3 ligand).			
GN	TNFRSF12 OR APO3L OR DR3LG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND N-TERMINUS OF SOLUBLE FORM.			
RC	TISSUE-Tonsil, and fetal liver;			
RC	MEDLINE=98070415; PubMed=9405449;			
RA	Chicheportiche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H., Hession C., Garcia I., Browning J.L.;			
RA	"TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis."			
RT	J. Biol. Chem. 272:32401-32410(1997).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE-Fetal kidney;			
RC	MEDLINE=98228355; PubMed=9560343;			
RA	Marsters S.A., Sheridan J.P., Pitti R.M., Brush J., Goddard A., Ashkenazi A.;			
RA	"Identification of a ligand for the death-domain-containing receptor Apo3."			
RT	Curr. Biol. 8:525-528(1998).			
RL	[3]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE-Tonsil;			
RC	Strausberg R.;			
RA	Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.			
RL	[4]			
RN	FUNCTION.			
RP	PubMed=10085077;			
RA	Lynch C.N., Wang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S.R.;			
RA	"TWEAK induces angiogenesis and proliferation of endothelial cells.;"			
RT	J. Biol. Chem. 274:8455-8459(1999).			
RL	J. Biol. Chem. 274:8455-8459(1999).			
CC	-1- FUNCTION: Binds to FN14 and possibly also to TNFRSF12/APO3. Weak inducer of apoptosis in some cell types. Mediates NF-kappaB activation. May promote angiogenesis and the proliferation of endothelial cells.			
CC	-1- SUBUNIT: HOMOTRIMER (POTENTIAL).			
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein and secreted.			
CC	-1- TISSUE SPECIFICITY: Highly expressed in adult heart, pancreas, skeletal muscle, brain, colon, small intestine, lung, ovary, prostate, spleen, lymph node, appendix and peripheral blood lymphocytes. Low expression in kidney, testis, liver, placenta, thymus and bone marrow. Also detected in fetal kidney, liver, lung and brain.			
CC	-1- PFM: The soluble form derives from the membrane form			

34	7	2.5	240	1	RNFE_PSEAE	09hyb5 pseudomonas
35	7	2.5	243	1	MOEB_HAETN	P45211 haemophilus
36	7	2.5	244	1	TNFC_HUMAN	O06643 homo sapien
37	7	2.5	249	1	MOEB_ECOLI	P12282 escherichia
38	7	2.5	249	1	MOEB_SALTY	O56067 salmonella
39	7	2.5	257	1	KDTX_SERVA	O54435 serratia ma
40	7	2.5	310	1	TNFC_MARMO	O91m10 marmota mon
41	7	2.5	316	1	ISPH_XYLEFA	O91m10 xylella fas
42	7	2.5	324	1	ODPB_BACSU	P21882 bacillus su
43	7	2.5	329	1	SRAB_CAEEL	O09208 caenorhabdi
44	7	2.5	335	1	LAG2_RAT	O35777 rattus norv
45	7	2.5	344	1	LEU3_THEAQ	P24098 thermus aqu

CC by proteolytic processing.  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -1- CAUTION: Ref.3 sequence differs from that shown due to a  
 CC frameshift in position 125.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF030099; AAC51923.1; -;  
 CC EMBL: AF055872; AAC39724.1; -;  
 CC EMBL: BC019047; AAH19047.1; ALT\_FRAME.  
 CC Gene: HGNC:11927; TNFSF12.  
 CC MIM: 602695; -  
 CC InterPro: IPR000478; TNF\_family.  
 CC Pfam: PF00229; TNF\_1.  
 CC SMART: SM00207; TNF\_1.  
 CC PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
 CC PROSITE: PS00049; TNF\_2; 1.  
 CC Cytokine: Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.  
 CC CHAIN 1 249  
 CC FT CHAIN 94 249  
 CC FT DOMAIN 1 21  
 CC FT TRANSMEM 22 42  
 CC FT DOMAIN 43 249  
 CC FT SITE 93 94  
 CC FT CARBOHYD 139 139  
 CC SQ SEQUENCE 249 AA; 27216 MW; E60843361C28BBA CRC64;  
 CC  
 CC Query Match 87.7%; Score 249; DB 1; Length 249;  
 CC Best Local Similarity 100.0%; Pred.No.3.2e-231;  
 CC Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC QY 36 MAARRSORRRGRGEPETALVPLALGLALACIGLLAVVSLGSRASLSAEPAGEEL 95  
 CC DB 1 MAARRSORRRGRGEPETALVPLALGLALACIGLLAVVSLGSRASLSAEPAGEEL 60  
 CC QY 96 VAEEDDPSELNQTREESQDPAPFLNLRVPRRSAPRGKTRARRAIAHYEYHPRGOD 155  
 CC DB 61 VAEEDDPSELNQTREESQDPAPFLNLRVPRRSAPRGKTRARRAIAHYEYHPRGOD 120  
 CC QY 156 GAGAGVDCVTSGWEARINSSPLRYNRQIGETVTRAGIYLYCQVHDEGKAVYIKLD 215  
 CC DB 121 GAGAGVDCVTSGWEARINSSPLRYNRQIGETVTRAGIYLYCQVHDEGKAVYIKLD 180  
 CC QY 216 LTVGVYALRCLREFSATASLSGLPOLRLCQVSGLLALRGSSLRITTLPMHAKAPFL 275  
 CC DB 181 LTVGVYALRCLREFSATASLSGLPOLRLCQVSGLLALRGSSLRITTLPMHAKAPFL 240  
 CC QY 276 TYFGLEQVH 284  
 CC DB 241 TYFGLEQVH 249  
 CC  
 CC RESULT 2  
 CC TN12\_MOUSE STANDARD; PRT; 225 AA.  
 CC ID TN12\_MOUSE  
 CC AC 054907; Q9CTP2;  
 CC DT 15-JUN-2002 (Rel. 41, Created)  
 CC DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 CC DE Tumor necrosis factor ligand superfamily member 12 (TNF-related weak  
 CC inducer of apoptosis) (TWEAK) (Fragment).  
 CC GN TNFSF12.  
 CC OS Mus musculus (Mouse).  
 CC EC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 CC NCBI\_TaxID=10090;  
 CC OX [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC TISSUE=Peritoneal macrophage;  
 CC RX MEDLINE=98070415; PubMed=9405449;  
 CC RA Chlcheporliche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,  
 CC RT Hesson C., Garcia I., Browning J.L.;  
 CC RT "TWEAK, a new secreted ligand in the tumor necrosis factor family that  
 CC weakly induces apoptosis."  
 CC RL J. Biol. Chem. 272:32401-32410(1997).  
 CC [2]  
 CC RP SEQUENCE OF 83-225 FROM N.A.  
 CC RC STRAIN=C57BL/6J; TISSUE=Retina;  
 CC RX MEDLINE=21085660; PubMed=11217851;  
 CC RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 CC RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 CC RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana T.,  
 CC RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 CC RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 CC RA Felschmann W., Gaasterland T., Glasl C., King B., Kochiwa H.,  
 CC RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 CC RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 CC RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 CC RA Blake J., Botfield D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 CC RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 CC RA Gwinnich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,  
 CC RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 CC RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 CC RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 CC RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,  
 CC RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,  
 CC RA Hayashizaki Y.;  
 CC RL "Functional annotation of a full-length mouse cDNA collection."  
 CC Nature 409:685-690(2001).  
 CC CC -1- FUNCTION: Binds to FN14 and possibly also to TNFRSF12/Ap03. Weak  
 CC inducer of apoptosis in some cell types. Promotes angiogenesis and  
 CC the proliferation of endothelial cells. Mediates NF-kappaB  
 CC activation (By similarity).  
 CC CC -1- SUBUNIT: Homotrimer (Potential).  
 CC CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By  
 CC similarity).  
 CC CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.  
 CC CC -1- PTM: The soluble form is produced from the membrane form by  
 CC proteolytic processing (By similarity).  
 CC CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF030100; AAC53517.1; -;  
 CC EMBL: AK020909; BAB32249.1; -;  
 CC MGD: MGI:1196259; Tnfsf12.  
 CC DR InterPro: IPR000478; TNF\_family.  
 CC DR Pfam: PF00229; TNF\_1.  
 CC DR SMART: SM00207; TNF\_1.  
 CC DR PROSITE: PS00251; TNF\_1; FALSE\_NEG.  
 CC DR PROSITE: PS00049; TNF\_2; 1.  
 CC KW Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.  
 CC FT NON\_TER 1 1  
 CC FT CHAIN <1 225  
 CC FT CHAIN 70 225  
 CC FT CHAIN 70 225  
 CC FT TRANSMEM <1 21  
 CC FT DOMAIN 22 225  
 CC FT SITE 69 70  
 CC  
 CC TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 CC MEMBER 12, MEMBRANE FORM.  
 CC TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY  
 CC MEMBER 12, SECRETED FORM (BY SIMILARITY).  
 CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 CC (POTENTIAL).  
 CC EXTRACELLULAR (POTENTIAL).  
 CC CLEAVAGE (BY SIMILARITY).

FT DISULFID 167 186 POTENTIAL.  
 FT CARBOHYD 115 115 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 225 AA; 24781 MW; 90C412CC0480659B CRC64;

Query Match 11.3%; Score 32; DB 1; Length 225;  
 Best Local Similarity 100.0%; Pred. No. 4.7e-23;  
 Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 RRAIAHYEHVPRGODGAGVGYSGWEE 170  
 |||||||  
 DB 80 RRAIAHYEHVPRGODGAGVGYSGWEE 111

## RESULT 3

YFBW\_ECOLI  
 ID YFBW\_ECOLI STANDARD; PRT; 111 AA.  
 AC Q47377;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein yfbw.  
 GN YFBW OR B2257.1.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_Taxid=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blatner F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [2]  
 RP SEQUENCE OF 81-111 FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=96186953; PubMed=8626063;  
 RA Sharma V., Hudepeth M.E., Meganathan R.;  
 RT "Menadiolone (vitamin K2) biosynthesis: localization and  
 RT characterization of the menE gene from Escherichia coli.";  
 RL Gene 168:43-48(1996).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: AE000315; NOT ANNOTATED\_CDS.  
 DR EMBL: L35031; BAB04895.1; -  
 DR Ecogene; EG14344; yfbw.  
 DR InterPro: IPR000620; DUF6.  
 DR Pfam; PF00892; DUF6; 1.  
 KW Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 36 56 POTENTIAL.  
 FT TRANSMEM 62 82 POTENTIAL.  
 FT TRANSMEM 88 108 POTENTIAL.  
 SQ SEQUENCE 111 AA; 12192 MW; 7CFA06D75DA33D69 CRC64;

Query Match 3.5%; Score 10; DB 1; Length 111;  
 Best Local Similarity 100.0%; Pred. No. 0.033;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGLALACIGL 72  
 |||||||  
 DB 40 LGLALACIGL 49

RESULT 4  
 PSAB\_ODOST STANDARD; PRT; 733 AA.  
 AC P49480;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Photosystem I P700 chlorophyll A apoprotein A2 (psab) (PST-B).  
 GN PSAB.  
 OS Odontella sinensis (Marine centric diatom).  
 OC Chloroplast.  
 OC Eukaryota; stramenopiles; Bacillariophyta; Coscinodiscophyceae;  
 OC Biddulphiophycidae; Eupodiscaceae; Eupodiscaceae; Odontella.  
 OX NCBI\_Taxid=2839;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kowallik K.V., Stoebe B., Schaffran I., Kroth-Pancic P., Freier U.;  
 RT "The chloroplast genome of a chlorophyll a+c-containing alga,  
 RT Odontella sinensis.";  
 RL Plant Mol. Biol. Rep. 13:336-342(1995).  
 CC -1- FUNCTION: Psab and psab bind P700, the primary electron donor of  
 CC photosystem I (PSI), as well as the electron acceptors A0, A1, and  
 CC FX. PSI functions as a plastocyanin/cytochrome c6-ferredoxin  
 CC oxidoreductase.  
 CC -1- COFACTOR: P700 is a chlorophyll A dimer, A0 is chlorophyll A, A1  
 CC is a phytylquinone and FX is a 4Fe-4S iron-sulfur center.  
 CC -1- SUBUNIT: A psab/B heterodimer binds the P700 chlorophyll special  
 CC pair and subsequent electron acceptors. The PSI reaction center of  
 CC higher plants and algae is composed of one at least 11 subunits.  
 CC  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast  
 CC thylakoid membrane.  
 CC  
 CC -1- SIMILARITY: BELONGS TO THE PSAA/PSAB FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: Z67753; CA91749.1; -  
 DR HSSP: P25897; LJB0.  
 DR InterPro: IPR001280; PSI\_Psaa/B.  
 DR Pfam; PF00223; psaa-psab; 1.  
 DR PRINTS; PR00257; PHOTOSPSAB.  
 DR PROSITE; PS00419; PHOTOSYSTEM\_I\_PSAAB; 1.  
 KW Photosynthesis; Photosystem I; Transport; Electron transport;  
 KW Chloroplast; Thylakoid; Transmembrane; Metal-binding; Iron;  
 KW Iron-sulfur; 4Fe-4S; Chlorophyll.  
 FT TRANSMEM 46 69  
 FT TRANSMEM 134 157  
 FT TRANSMEM 174 198  
 FT TRANSMEM 272 290  
 FT TRANSMEM 329 352  
 FT TRANSMEM 368 394  
 FT TRANSMEM 416 438  
 FT TRANSMEM 516 534  
 FT TRANSMEM 574 595  
 FT TRANSMEM 642 664  
 FT TRANSMEM 706 726  
 FT TRANSMEM 766 786  
 FT METAL 558 558  
 FT METAL 567 567  
 FT BINDING 653 653

FT BINDING 661 661  
 FT BINDING 669 669  
 FT BINDING 670 670  
 FT BINDING 699 699  
 SQ SEQUENCE 733 AA; 82103 MW; 13439AF1E441BB7 CRC64;

Query Match

3.2%; Score 9; DB 1; Length 733;

Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 LGIACLG 71  
| | | | | | | | | |  
Db 333 LGIACLG 341

RESULT 5  
PSAB\_CVACA STANDARD; PRT; 734 AA.  
ID PSAB\_CVACA  
AC 09106;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Photosystem I P700 chlorophyll A apoprotein A2 (PsaB) (PSI-B).  
GN PSAB.  
OS Cyanidium caldarium.  
OC Chloroplast.  
OC Eukaryota; Rhodophyta; Porphyridiales; Porphyridiaceae;  
OC Cyanidium.  
OX NCBI\_TaxID=2771;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RK-1;  
RX MEDLINE=20496959; PubMed=11040290;  
RA Gloeckner G., Rosenthal A., Valentin K.-U.;  
RT "The structure and gene repertoire of an ancient red algal plastid genome."  
RT J. Mol. Evol. 51:382-390(2000).  
CC -1- FUNCTION: PsaA and PsaB bind P700, the primary electron donor of photosystem I (PSI), as well as the electron acceptors A0, A1, and FX. PSI functions as a plastocyanin/cytochrome c6-ferredoxin oxidoreductase.  
CC -1- COFACTOR: P700 is a chlorophyll A dimer, A0 is chlorophyll A, A1 is a phytylquinone and FX is a 4Fe-4S iron-sulfur center.  
CC -1- SUBUNIT: A psaA/B heterodimer binds the P700 chlorophyll special pair and subsequent electron acceptors. The PSI reaction center of higher plants and algae is composed of one at least 11 subunits.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast thylakoid membrane.  
CC -1- SIMILARITY: BELONGS TO THE PSA/PSAB FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AF022186; AAF12881.1; -  
CC HSSP: P25897; LUBO.  
CC InterPro: IPR001280; PSI\_PsaA/B.  
CC Pfam: PF00223; PsaA\_PsaB.1  
CC PRINTS: PR00257; PHOTOSPSAB.  
CC PROSITE: PS00419; PHOTOSYSTEM\_I\_PSAAB.1.  
CC Photosynthesis; Photosystem I; Transport; Electron transport;  
CC Chloroplast; Thylakoid; Transmembrane; Metal-binding; Iron;  
CC Iron-sulfur; 4Fe-4S; Chlorophyll.  
CC TRANSMEM 46 69 I (POTENTIAL).  
CC TRANSMEM 135 158 II (POTENTIAL).  
CC TRANSMEM 175 199 III (POTENTIAL).  
CC TRANSMEM 273 291 IV (POTENTIAL).  
CC TRANSMEM 330 353 V (POTENTIAL).  
CC TRANSMEM 369 395 VI (POTENTIAL).  
CC TRANSMEM 417 439 VII (POTENTIAL).  
CC TRANSMEM 517 533 VIII (POTENTIAL).  
CC TRANSMEM 575 596 IX (POTENTIAL).  
CC TRANSMEM 643 665 X (POTENTIAL).  
CC TRANSMEM 707 727 XI (POTENTIAL).  
CC TRANSMEM 559 559 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).  
CC TRANSMEM 568 568 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).  
CC METAL

FT BINDING 654 654 P700 SPECIAL PAIR CHLOROPHYLL AXIAL  
FT BINDING 662 662 LIGAND (BY SIMILARITY).  
FT BINDING 670 670 A0 CHLOROPHYLL (BY SIMILARITY).  
FT BINDING 671 671 A1 CHLOROPHYLL (BY SIMILARITY).  
FT BINDING 700 700 A1 PHYTYLQUINONE (BY SIMILARITY).  
FT BINDING 700 700 A1 PHYTYLQUINONE (BY SIMILARITY).  
SQ SEQUENCE 734 AA; 82359 MW; 4496AA2AE39CA9B9 CRC64;

Query Match 3.2%; Score 9; DB 1; Length 734;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 LGIACLG 71  
| | | | | | | | | |  
Db 334 LGIACLG 342

RESULT 6  
ADHS\_GLUOX STANDARD; PRT; 179 AA.  
ID ADHS\_GLUOX  
AC 005544;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Alcohol dehydrogenase 15 kDa subunit precursor (G3-ADH subunit III).  
GN ADHS.  
OS Gluconobacter oxydans (Gluconobacter suboxydans).  
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;  
OC Gluconobacter.  
OX NCBI\_TaxID=442;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-40.  
RC STRAIN-IFO 12528;  
RX MEDLINE=97208225; PubMed=9055427;  
RA Kondo K., Horinouchi S.;  
RT "Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Gluconobacter suboxydans and their expression in *Acetobacter pasteurianus*."  
RT Appl. Environ. Microbiol. 63:1131-1138(1997).  
CC -1- FUNCTION: NOT ESSENTIAL FOR ALCOHOL DEHYDROGENASE ACTIVITY.  
CC -1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.  
CC ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.  
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: D86440; BAA19756.1; -  
CC Membrane; Periplasmic; Signal.  
CC SIGNAL 1 24 POTENTIAL.  
CC CHAIN 25 179 ALCOHOL DEHYDROGENASE 15 KDA SUBUNIT.  
FT MOD.RES 25 25 PYRROLIDONE CARBOXYLIC ACID.  
SQ SEQUENCE 179 AA; 19943 MW; F6AF2436563C66 CRC64;  
Query Match 2.8%; Score 8; DB 1; Length 179;  
Best Local Similarity 100.0%; Pred. No. 4.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 LAIGIGLA 66  
| | | | | | | | | |  
Db 11 LAIGIGLA 18

RESULT 7  
Y304\_BROME STANDARD; PRT; 220 AA.  
ID Y304\_BROME  
AC 08YD73;



```
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein BME110304.
GN BME110304.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756588;
RA Delvecchio V.G., Kapatal V., Redkar R.J., Patra G., Mujar C., los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonksi L., Larsen N., D'Souza M., Bernal A., Golsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Lelesson J.-J.,
RA Hesselkorn R., Kyriades N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis."
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE UPF0191 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: AEO09669; AAL53546.1;
KM Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 20 39 POTENTIAL.
FT TRANSMEM 54 72 POTENTIAL.
FT TRANSMEM 85 104 POTENTIAL.
FT TRANSMEM 124 146 POTENTIAL.
FT TRANSMEM 153 175 POTENTIAL.
FT TRANSMEM 179 198 POTENTIAL.
SQ SEQUENCE 220 AA; 24815 MW; 182C0244743B17FA CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 220;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 ALVPLAL 61
Db 132 ALVPLAL 139
|||||
|
```

```
CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
CC ADENYLATE CYCLASE (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC -----
DR EMBL: AJ245705; CAB53398.1;
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE: PS50262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 37 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 38 63 1 (POTENTIAL).
FT DOMAIN 64 72 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 73 93 2 (POTENTIAL).
FT DOMAIN 94 118 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 119 140 3 (POTENTIAL).
FT DOMAIN 141 163 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 164 183 4 (POTENTIAL).
FT DOMAIN 184 191 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 192 211 5 (POTENTIAL).
FT DOMAIN 212 240 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 241 266 6 (POTENTIAL).
FT DOMAIN 267 279 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 280 300 7 (POTENTIAL).
FT DOMAIN 301 317 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 315 315 PALMITATE (POTENTIAL).
SQ SEQUENCE 317 AA; 34699 MW; 6615D2146E1D247F CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 1; Length 317;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 RARRAIA 144
Db 160 RARRAIA 167
|||||
|
```

```
RESULT 8
MSHR_PANTR STANDARD; PRT; 317 AA.
AC 09JUK4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Melanocyte stimulating hormone receptor (MSH-R) (Melanotropin
DE receptor) (Melanocortin-1 receptor) (MC1-R).
GN MC1R.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Rees J.L., Harding R.M., Healy E., Jackson I.J., Ray A.J., Ellis N.S.,
RA Flanagan N., Todd C., Dixon C., Matthews J.N., Sanjanta A.,
RA Birch-Machin M.A.;
RT "Chimpanzee melanocortin 1 sequence."
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: RECEPTOR FOR MSH (ALPHA, BETA AND GAMMA) AND ACTH. THE
```

```
RESULT 9
FDRB_METUA STANDARD; PRT; 379 AA.
ID FDRB_METUA
AC Q60316;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative formate dehydrogenase beta chain (EC 1.2.1.2).
GN MD0005.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerevage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek K., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrman J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
```

RA Klenk H.-P., Frazer C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*  
 RL *jannaschii*.";  
 CC Science 273:1058-1073(1996).  
 CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) -> CO(2) + NADH.  
 CC -1- COFACTOR: BINDS TWO 4FE-4S CLUSTERS (PROBABLY).  
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS (BY SIMILARITY).  
 CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF  
 CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.  
 CC -1- SIMILARITY: STRONG, TO THE BETA SUBUNIT OF *M.THERMOAUTOTROPHICUM*  
 CC FDH.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: U67459; AAB97986.1; -  
 CC TRIGR: M00005; -  
 CC InterPro: IPR001450; 4Fe4S\_ferredoxin.  
 CC Pfam: PF000037; fer4; 1.  
 CC DR PROSITE: PS00198; 4FE4S\_FERREDOXIN, 2.  
 CC DR Hypothetical protein; Oxidoreductase; NAD; Electron transport;  
 CC KM Iron-sulfur; 4Fe-4S; Complete proteome.  
 CC FT METAL 280 280 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 283 283 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 286 286 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 290 290 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 330 330 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 333 333 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 336 336 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 340 340 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC SO SEQUENCE 379 AA; 43014 MW; 9C257CCAD5547F5A CRC64;  
 CC -----  
 CC Query Match 2.8%; Score 8; DB 1; Length 379;  
 CC Best Local Similarity 100.0%; Pred. No. 7.8;  
 CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC -----  
 CC 217 LVDGVIAL 224  
 CC Db 35 LVDGVIAL 42  
 CC -----  
 CC RESULT 10  
 CC ZP3\_MESAU STANDARD: PRT; 422 AA.  
 CC AC P23491;  
 CC DT 01-NOV-1991 (Rel. 20, Created)  
 CC DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 CC DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 CC DE zona pellucida sperm-binding protein 3 precursor (zona pellucida  
 CC DE glycoprotein zp3) (Sperm receptor) (zona pellucida protein C).  
 CC GN ZP3.  
 CC OS Mesocricetus auratus (Golden hamster).  
 CC OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 CC OC Mesocricetus.  
 CC OX NCBI\_TaxID=10036;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC TISSUE=Ovary;  
 CC RX MEDLINE=91078540; PubMed=2257975;  
 CC RA Klenk H.-P., Frazer C.M., Smith H.O., Woese C.R., Venter J.C.;  
 CC RT "Complete genome sequence of the methanogenic archaeon, *Methanococcus*  
 CC RL *jannaschii*.";  
 CC CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) -> CO(2) + NADH.  
 CC -1- COFACTOR: BINDS TWO 4FE-4S CLUSTERS (PROBABLY).  
 CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS (BY SIMILARITY).  
 CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF  
 CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.  
 CC -1- SIMILARITY: STRONG, TO THE BETA SUBUNIT OF *M.THERMOAUTOTROPHICUM*  
 CC FDH.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: U67459; AAB97986.1; -  
 CC TRIGR: M00005; -  
 CC InterPro: IPR001450; 4Fe4S\_ferredoxin.  
 CC Pfam: PF000037; fer4; 1.  
 CC DR PROSITE: PS00198; 4FE4S\_FERREDOXIN, 2.  
 CC DR Hypothetical protein; Oxidoreductase; NAD; Electron transport;  
 CC KM Iron-sulfur; 4Fe-4S; Complete proteome.  
 CC FT METAL 280 280 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 283 283 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 286 286 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 290 290 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 330 330 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 333 333 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 336 336 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC FT METAL 340 340 IRON-SULFUR (4FE-4S) (BY SIMILARITY).  
 CC SO SEQUENCE 379 AA; 43014 MW; 9C257CCAD5547F5A CRC64;

CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Extracellular  
 CC matrix.  
 CC -1- TISSUE SPECIFICITY: OOCYTES.  
 CC -1- DEVELOPMENTAL STAGE: GROWING OOCYTES.  
 CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.  
 CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: M63629; AAA37079.1; -  
 CC InterPro: IPR001507; Endoglin/CD105.  
 CC Pfam: PF00100; zona-pellucida; 1.  
 CC DR PRINTS: PR00023; ZPPELLUCIDA.  
 CC DR SMART: SM00241; ZP; 1.  
 CC DR PROSITE: PS00682; ZP; 1.  
 CC DR Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;  
 CC KM Extracellular matrix.  
 CC FT SIGNAL 1 22  
 CC FT CHAIN 23 422  
 CC FT DOMAIN 23 386  
 CC FT TRANSMEM 387 407  
 CC FT DOMAIN 408 422  
 CC FT DOMAIN 45 306  
 CC FT DOMAIN 119 158  
 CC FT DOMAIN 208 257  
 CC FT CARBOHYD 146 146  
 CC FT CARBOHYD 271 271  
 CC FT CARBOHYD 302 302  
 CC SO SEQUENCE 422 AA; 45827 MW; D0F95B57FE87EB01 CRC64;  
 CC -----  
 CC Query Match 2.8%; Score 8; DB 1; Length 422;  
 CC Best Local Similarity 100.0%; Pred. No. 8.6;  
 CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 CC -----  
 CC 59 LALGIGLA 66  
 CC Db 386 LALGIGLA 393  
 CC -----  
 CC RESULT 11  
 CC CYDC\_HAERIN STANDARD: PRT; 576 AA.  
 CC AC P45081;  
 CC DT 01-NOV-1995 (Rel. 32, Created)  
 CC DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 CC DE Transport ATP-binding protein cydc.  
 CC GN CYDC OR H1156.  
 CC OS Haemophilus influenzae.  
 CC OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 CC OC Haemophilus.  
 CC OX NCBI\_TaxID=727;  
 CC RN [1]  
 CC RP SEQUENCE FROM N.A.  
 CC RC STRAIN=RD / KW20 / ATCC 51907;  
 CC RX MEDLINE=95350630; PubMed=7542800;  
 CC RA Keriavane A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 CC RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Goeyne J.D.,  
 CC RA Scott J.D., Shirley R., Liu L.-I., Glodet A., Kelley J.M.,  
 CC RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 CC RA Uteirbeck T.R., Hanna M.C., Nguyen D.T., Saudak D.M., Brandon R.C.,  
 CC RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,  
 CC RA Gnehm C.L., McDonald L.A., Small K.V., Frazer C.M., Smith H.O.,  
 CC RA Venter J.C.;  
 CC RT "Whole-genome random sequencing and assembly of *Haemophilus influenzae*

Rd.":  
RL Science 269:496-512(1995).  
CC -1- FUNCTION: SOMEHOW INVOLVED IN THE CYTOCHROME D BRANCH OF AEROBIC  
CC RESPIRATION. SEEMS TO BE A COMPONENT OF A TRANSPORT SYSTEM  
CC (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Probable).  
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MSBA SUBFAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: U32795; AAC22811.1; -  
DR TIGR: H1156; -  
DR InterPro: IPR003593; AAA\_ATPase.  
DR InterPro: IPR003439; ABC\_transport.  
DR InterPro: IPR01140; ABCtransport.  
DR Pfam: PF00005; ABC\_tran; 1.  
DR Pfam: PF00664; ABC\_membrane; 1.  
DR ProDom: PD000006; ABC\_transport; 1.  
DR SMART: SM00382; AAA; 1.  
DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
DR ATP-binding; Transport; Transmembrane; Inner membrane;  
KW Complete proteome.  
FT TRANSMEM 16  
FT TRANSMEM 36  
FT TRANSMEM 58  
FT TRANSMEM 133  
FT TRANSMEM 155  
FT TRANSMEM 175  
FT TRANSMEM 244  
FT TRANSMEM 264  
FT TRANSMEM 281  
FT TRANSMEM 301  
FT NP\_BIND 372  
FT NP\_BIND 379  
SQ SEQUENCE 576 AA; 64831 MW; A9ACD8B9B294B1B3 CRC64;  
  
Query Match 2.8%; Score 8; DB 1; Length 576;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 58 PLALGIGL 65  
Db 159 PLALGIGL 166  
  
RESULT 12  
GGTS\_HUMAN  
ID GGTS\_HUMAN STANDARD; PRT; 586 AA.  
AC P36269; O96PCL; -  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Gamma-glutamyltranspeptidase 5 precursor (EC.2.3.2.2) (Gamma-  
DE glutamyltransferase 5) (GGR-Rel).  
GN GGTAL OR GGTS.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RP [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Placenta;  
RA MEDLINE-91296809; Pubmed-1676842;  
RA Heisterkamp N., Rajpert-De Meyts E., Uribe L., Forman H.J.,  
RA Groffen J.;  
RT "Identification of a human gamma-glutamyl cleaving enzyme related to,  
RT but distinct from, gamma-glutamyl transpeptidase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:6303-6307(1991).  
RN [2]  
RN SEQUENCE FROM N.A.  
RP TISSUE-Brain;

RA Strausberg R.;  
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.  
CC -1- FUNCTION: CAN HYDROLYZE THE GAMMA-GLUTAMYL MOIETY OF GLUTATHIONE;  
CC AS WELL AS CONVERT LEUKOTRIENE C4 TO LEUKOTRIENE D4.  
CC -1- CATALYTIC ACTIVITY: (5-L-glutamyl)-peptide + an amino acid -  
CC peptide + 5-L-glutamyl-amino acid.  
CC -1- SUBUNIT: THIS ENZYME CONSISTS OF TWO POLYPEPTIDE CHAINS, WHICH ARE  
CC SYNTHESIZED IN PRECURSOR FORM FROM A SINGLE POLYPEPTIDE (BY  
CC SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Probable).  
CC -1- SIMILARITY: STRONG, TO OTHER GGT ENZYMES AND TO GL-7ACA ACYLASES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
DR EMBL: M64099; AAS58503.1; -  
DR EMBL: BC011362; AAH11362.1; -  
DR PIR: A41125; A41125.  
DR MEROPS: T03.002; -  
DR GeneW: HGNC:4260; GGTAL1.  
DR MIM: 137168; -  
DR InterPro: IPR000101; Gglutinsptidase.  
DR Pfam: PF01019; G-glu\_transpept; 1.  
DR PRINTS: PR01210; GGTTRANSPASE.  
DR PROSITE: PS00462; G-GLU\_TRANSPEPTIDASE; 1.  
DR Transferrase; Acyltransferase; Zymogen; Glycoprotein; Transmembrane;  
KW Glutathione biosynthesis; Signal-anchor.  
FT CHAIN 1  
FT CHAIN 387  
FT CHAIN 388  
FT CHAIN 586  
FT TRANSMEM 7  
FT TRANSMEM 29  
FT DOMAIN 30  
FT CARBOHYD 98  
FT CARBOHYD 98  
FT CARBOHYD 204  
FT CARBOHYD 204  
FT CARBOHYD 303  
FT CARBOHYD 303  
FT CARBOHYD 347  
FT CARBOHYD 347  
FT CARBOHYD 535  
FT CARBOHYD 535  
FT CARBOHYD 550  
FT CARBOHYD 550  
FT CONFLICT 330  
FT CONFLICT 437  
SQ SEQUENCE 586 AA; 62319 MW; 1B5543CB0934B16B CRC64;  
  
Query Match 2.8%; Score 8; DB 1; Length 586;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 61 LGIGLALA 68  
Db 14 LGIGLALA 21  
  
RESULT 13  
HSP1\_SAGIM  
ID HSP1\_SAGIM STANDARD; PRT; 49 AA.  
AC P24714; -  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Sperm protamine P1 (Cysteine-rich protamine).  
GN PRM1.  
OS Saguinus imperator (Tamarin).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.  
OX NCBI\_TaxID=9491;  
RP [1]  
RP SEQUENCE FROM N.A.

RC TISSUE-liver;  
 RA MEDLINE-92051332; PubMed-1840669;  
 RX Queralt R., Oliva R.;  
 RT "Protamine 1 gene sequence from the primate *Saguinus imperator*  
 RL isolated with PCR using consensus oligonucleotides.";  
 CC Nucleic Acids Res. 19:5786-5786(1991).  
 CC -1- FUNCTION: PROTAHINS SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF  
 CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT  
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.  
 CC -1- SUBUNIT: CROSS-LINKED BY INTERCHAIN DISULFIDE BONDS AROUND THE  
 CC DNA-HELIX (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: TESTIS.  
 CC -1- SIMILARITY: BELONGS TO THE PROTAHINE P1 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X61678; CAA3853.1; -;  
 DR PIR: S22582; S22582.  
 DR InterPro: IPR000221; Protamine\_p1.  
 DR Pfam: PF00260; Protamine\_p1; 1.  
 DR PROSITE: PS00048; PROTAHINE\_P1; 1.  
 KM Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;  
 KW Testis; DNA condensation; Nuclear protein.  
 FT INIT\_MET 0  
 FT SEQUENCE 49 AA; 6545 MW; 8399C403F5B207F6 CRC64;  
 SQ  
 Query Match 2.5%; Score 7; DB 1; Length 49;  
 Best Local Similarity 100.0%; Pred. No. 12;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 ORRRGR 48  
 DB 17 ORRRGR 23

RESULT 14  
 HSP1\_DIDMA STANDARD; PRT; 57 AA.  
 AC P35305;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Sperm protamine P1.  
 GN PPM1.  
 OS *Monodelphis domestica* (Short-tailed grey opossum), and  
 OS *Monodelphis domestica* (Short-tailed grey opossum).  
 OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Eutelestomi;  
 OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.  
 OC NCBI\_TaxID=9267; 13616;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-D. marsupialis;  
 RX MEDLINE-93345500; PubMed-8344286;  
 RA Winkfein R.J., Nishikawa S., Connor W., Dixon G.H.;  
 RT "Characterization of a marsupial sperm protamine gene and its  
 RT transcripts from the North American opossum (*Didelphis*  
 RT *marsupialis*).";  
 RL Eur. J. Biochem. 215:63-72(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-M. domestica;  
 RX MEDLINE-95215351; PubMed-7700877;  
 RA Redif J.D., Krajewski C., Western M., Winkfein R.J., Dixon G.H.;  
 RT "Molecular phylogeny and evolution of marsupial protamine P1 genes.";  
 RL Proc. R. Soc. Lond. B, Biol. Sci. 259:7-14(1993).  
 CC -1- FUNCTION: PROTAHINS SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF

CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT  
 CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: TESTIS.  
 CC -1- SIMILARITY: BELONGS TO THE PROTAHINE P1 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L17007; AAA02812.1; -;  
 DR EMBL: X74044; CAA52193.1; -;  
 DR EMBL: L35448; AAA74612.1; -;  
 DR PIR: S34045; S34045.  
 DR InterPro: IPR000221; Protamine\_p1.  
 DR Pfam: PF00260; Protamine\_p1; 1.  
 DR PROSITE: PS00048; PROTAHINE\_P1; 1.  
 KM Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;  
 KW Testis; DNA condensation; Nuclear protein.  
 FT INIT\_MET 0  
 FT SEQUENCE 57 AA; 7810 MW; 283715B280214E52 CRC64;  
 SQ  
 Query Match 2.5%; Score 7; DB 1; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRG 49  
 DB 34 RRRGRG 40

RESULT 15  
 A62F\_DROME STANDARD; PRT; 115 AA.  
 AC Q46202;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Accessory gland protein Acp62F precursor.  
 GN ACP62F OR CG1262.  
 GN Drosophila melanogaster (Fruit fly).  
 OS Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscophora; Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.; FUNCTION, AND TISSUE SPECIFICITY.  
 RC STRAIN-Canton-S; TISSUE-Male accessory gland;  
 RX MEDLINE-98135120; PubMed-9474779;  
 RA Wolfner M.F., Harada H.A., Bertram M.J., Stelick T.J., Kraus K.W.,  
 RA Kalb J.M., Lung Y.O., Neudam D.M., Park M., Trim V.K.;  
 RT "New genes for male accessory gland proteins in *Drosophila*  
 RT *melanogaster*.";  
 RL Insect Biochem. Mol. Biol. 27:825-834(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Berkeley;  
 RX MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
 RA Bailes R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brottler P.,  
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,

Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaser K.,  
 Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwan C.,  
 Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,  
 Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 Palazzolo M., Pitman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 "The genome sequence of *Drosophila melanogaster*.";  
 Science 287:2185-2195(2000).  
 [3]  
 SEQUENCE OF 7-111 FROM N.A.  
 RC STRAIN-Various strains;  
 RX MEDLINE-20556153; Pubmed-11102381;  
 RA Begun D.J., Whitely P., Todd B.L., Waldrif-Dall H.M., Clark A.G.;  
 RT "Molecular population genetics of male accessory gland proteins in  
*Drosophila*.";  
 RL Genetics 156:1879-1888(2000).  
 CC -1- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGES IN  
 MATED FEMALE FLIES. MAY CONTRIBUTE TO THE TOXICITY OF SEMINAL  
 FLUID AND THE DECREASED LIFE-SPAN OF MATED FEMALES. MAY ALSO  
 AFFECT NEUROMUSCULAR EVENTS AFTER MATING CONCERNING SPERM STORAGE  
 AND ECG RELEASE.  
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).  
 CC -1- TISSUE SPECIFICITY: SEMINAL FLUID.  
 CC -1- SIMILARITY: SOME, TO P. NIGRIVENTER TX2-6.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; 085763; AAB96387.1; -  
 DR EMBL; AE003475; AAF47683.1; -  
 DR EMBL; AY010608; AAG35367.1; -  
 DR EMBL; AY010609; AAG35368.1; -  
 DR EMBL; AY010610; AAG35369.1; -  
 DR EMBL; AY010611; AAG35370.1; -  
 DR EMBL; AY010612; AAG35371.1; -  
 DR EMBL; AY010613; AAG35372.1; -  
 DR EMBL; AY010614; AAG35373.1; -  
 DR EMBL; AY010615; AAG35374.1; -  
 DR EMBL; AY010616; AAG35375.1; -  
 DR EMBL; AY010617; AAG35376.1; -  
 DR FlyBase; FBgn0020509; AcP62F.  
 DR InterPro; IPR002919; TIL\_Cysrich.  
 DR Pfam; PF01826; TIL; 1.  
 KW Behavior; Signal.  
 FT SIGNAL 1 24 POTENTIAL.  
 CHAIN 25 115 ACCESSORY GLAND PROTEIN ACPE62F.  
 SEQUENCE 115 AA; 12570 MW; 4326AA6FC32291D CRC64;

Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 68 AGLILL 74  
 Db 11 AGLILL 17  
 RESULT 16  
 Y151\_ECOLI STANDARD; PRT; 118 AA.  
 AC P03838;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Insertion element IS5 very hypothetical 12 kDa protein.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-82028653; Pubmed-6269959;  
 RA Schoner B., Kahn M.;  
 RT "The nucleotide sequence of IS5 from *Escherichia coli*.";  
 RL Gene 14:165-174(1981).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-82028652; Pubmed-6269958;  
 RA Engler J.A., van Bree M.P.;  
 RT "The nucleotide sequence and protein-coding capability of the  
 transposable element IS5.";  
 RL Gene 14:155-163(1981).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,  
 RA Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurd O.,  
 RA Laskari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,  
 RA Davis R.W.;  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,  
 RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R.,  
 RA Ichihara S., Mizuno T., Makino K., Nakata A., Yura T., Sempel G.,  
 RA Mitsuuchi K.;  
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE-97251357; Pubmed-9097039;  
 RA Itoh T., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,  
 RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,  
 RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,  
 RA Motomura K., Nakano S., Nakamura Y., Nishimoto H., Nishio Y.,  
 RA Oshima T., Saito N., Sempel G., Seki Y., Sivasubram S.,  
 RA Yamamoto Y., Takeda J., Takemoto K., Takeuchi Y., Wada C.,  
 RA Yamamoto Y., Horiuchi T.;  
 RT "A 570-kb DNA sequence of the *Escherichia coli* K-12 genome  
 corresponding to the 28.0-40.1 min region on the linkage map.";  
 RL DNA Res. 3:363-377(1996).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; J01734; -; NOT ANNOTATED\_CDS.  
 DR EMBL; U70214; AAB08680.1; -

Query Match 2.5%; Score 7; DB 1; Length 115;

DR EMBL: D83536; -: NOT\_ANNOTATED\_CDS.  
 DR EMBL: D90771; BAA14925.1; -  
 DR EMBL: D90772; BAA14935.1; -  
 DR EMBL: D90831; BAA15715.1; -  
 DR EMBL: D90841; BAA15872.1; -  
 DR EMBL: D90847; BAA15958.1; -  
 DR EMBL: D90848; BAA15963.1; -  
 DR PIR: A04466; IECSB.  
 KW Hypothetical protein; Transposable element.  
 SQ SEQUENCE 118 AA; 12270 MW; 348014FAC765058E CRC64;

Query Match 2.5%; Score 7; DB 1; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 26;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 23 DCGAVRQ 29  
 11111111  
 Db 107 DCGAVRQ 113

RESULT 17  
 IL13\_MOUSE STANDARD; PRT; 131 AA.  
 AC P20109;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Interleukin-13 precursor (IL-13) (T-cell activation protein p600).  
 GN IL13 OR IL-13.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89093958; PubMed=2521353;  
 RA Brown K.D., Zurewaki S.M., Mosmann T.R., Zurewaki G.;  
 RT "A family of small inducible proteins secreted by leukocytes are  
 RT members of a new superfamily that includes leukocyte and  
 RT fibroblast-derived inflammatory agents, growth factors, and  
 RT indicators of various activation processes.";  
 RL J. Immunol. 142:679-687(1989).  
 CC -1- FUNCTION: CYTOKINE. INHIBITS INFLAMMATORY CYTOKINE PRODUCTION.  
 CC SYNERGIZES WITH IL2 IN REGULATING INTERFERON-GAMMA SYNTHESIS.  
 CC MAY BE CRITICAL IN REGULATING INTERFERON-GAMMA SYNTHESIS.  
 CC (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO THE IL-4 / IL-13 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M23504; AAA40149.1; -  
 DR PIR: E30552; E30552.  
 DR HSSP: P35225; 31TR.  
 DR MGD: MGI:96541; 1113.  
 DR InterPro: IPR003634; Interleukin\_13.  
 DR InterPro: IPR001325; Interleukin\_4\_13.  
 DR Pfam: PF03487; Interleukin\_13; 1.  
 DR ProDom: PD015987; Interleukin\_13; 1.  
 DR SMART: SM00190; IL4\_13; 1.  
 DR PROSITE: PS00838; INTERLEUKIN\_4\_13; 1.  
 KW Cytokine; Glycoprotein; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 1 21 BY SIMILARITY.  
 FT DISULFID 22 131 INTERLEUKIN-13.  
 FT DISULFID 51 79 BY SIMILARITY.  
 FT CARBOHYD 67 93 BY SIMILARITY.  
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 52 52 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 75 75 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 131 AA; 14107 MW; 954F93F105713FED CRC64;

Query Match 2.5%; Score 7; DB 1; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 LALACLG 71  
 11111111  
 Db 9 LALACLG 15

RESULT 18  
 IL13\_RAT STANDARD; PRT; 131 AA.  
 ID IL13\_RAT  
 AC P42203;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Interleukin-13 precursor (IL-13) (T-cell activation protein p600).  
 GN IL13 OR IL-13.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=Kidney cortex;  
 RX MEDLINE=94092158; PubMed=7916615;  
 RA Lakkis F.G., Cruet E.N.;  
 RT "Cloning of rat Interleukin-13 (IL-13) cDNA and analysis of IL-13  
 RT gene expression in experimental glomerulonephritis.";  
 RL Biochem. Biophys. Res. Commun. 197:612-618(1993).  
 CC -1- FUNCTION: CYTOKINE. INHIBITS INFLAMMATORY CYTOKINE PRODUCTION.  
 CC SYNERGIZES WITH IL2 IN REGULATING INTERFERON-GAMMA SYNTHESIS.  
 CC MAY BE CRITICAL IN REGULATING INTERFERON-GAMMA SYNTHESIS.  
 CC (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- SIMILARITY: BELONGS TO THE IL-4 / IL-13 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L26913; AAA16478.1; -  
 DR HSSP: P35225; 31TR.  
 DR InterPro: IPR003634; Interleukin\_13.  
 DR InterPro: IPR001325; Interleukin\_4\_13.  
 DR Pfam: PF03487; Interleukin\_13; 1.  
 DR ProDom: PD015987; Interleukin\_13; 1.  
 DR SMART: SM00190; IL4\_13; 1.  
 DR PROSITE: PS00838; INTERLEUKIN\_4\_13; 1.  
 KW Cytokine; Glycoprotein; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 1 21 BY SIMILARITY.  
 FT DISULFID 22 131 INTERLEUKIN-13.  
 FT DISULFID 52 80 BY SIMILARITY.  
 FT CARBOHYD 68 94 BY SIMILARITY.  
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 76 76 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 131 AA; 14093 MW; E5008CAB8DE8C201 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 LALACLG 71

DB 9 LALACIG 15

RESULT 19  
YK01\_PYRHO STANDARD: PRT; 147 AA.

AC 05781;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein PH2001.  
GN PH2001.

OS Pyrococcus horikoshii.  
OC Archaeae; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.

OK NCBI\_TaxID=53953;  
RN (1)  
RP SEQUENCE FROM N.A.

RX MEDLINE=96844137; PubMed=9679194;  
RC STRAIN=OT3;

RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,  
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,  
RA Masuchi Y., Shizuya H., Kikuchi H.;  
RT "Complete sequence and gene organization of the genome of a hyper-  
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res. 5:55-76(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: AP000001; BAA31940.1; -;  
DR EMBL: AP000007; BAA31940.1; JOINED.  
DR EMBL: AP000007; BAA31943.1; -;  
DR EMBL: AP000001; BAA31943.1; JOINED.  
KW Hypothetical protein; Transmembrane; Complete proteome.  
FT TRANSMEM 41 61 POTENTIAL.  
FT TRANSME 67 87 POTENTIAL.  
SQ SEQUENCE 147 AA; 15324 MW; 247ED12FCEFE265B9 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 147;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 LALCIGL 73  
DB 45 LALCIGL 51

RESULT 20  
TNFC\_PIG STANDARD: PRT; 150 AA.  
AC 097SV8;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lymphotoxin-beta (Lr-beta) (Tumor necrosis factor C) (TNF-C) (Tumor  
DE necrosis factor ligand superfamily member 3) (Fragment).  
GN LTB OR TNFSF3 OR TNFC.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euteria; Cetartiodactyla; Suina; Suidae; Sus.  
OK NCBI\_TaxID=9823;  
RN (1)  
RP SEQUENCE FROM N.A.

RC STRAIN=Large white; TISSUE=Fibroblast;

RX MEDLINE=21108615; PubMed=11169259;

RA Chardon P., Rogel-Galliard C., Cattolico L., Duprat S., Vaiman M.,

RA Renard C.;

RT "Sequence of the swine major histocompatibility complex region

RT containing all non-classical class I genes.";

RL Tissue Antigens 57:55-65(2001).

CC -1- FUNCTION: Cytokine that binds to LTB $\beta$ /TNFRSP3. May play a specific  
CC role in immune response regulation. Provides the membrane anchor  
CC for the attachment of the heterotrimeric complex to the cell  
CC surface.

CC -1- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits or  
CC (less prevalent) two LTB and one LTB subunits.

CC -1- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).

CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: AJ251914; CAB63851.1; -;

DR InterPro: IPR003636; TNF\_abc.

DR InterPro: IPR004478; TNF\_family.

DR Pfam: PF00229; TNF; 1.

DR PRINTS: PR01234; TNECROSISFCT.

DR PRODOM: PD002012; TNF\_abc; 1.

DR SMART: SM00207; TNF; 1.

DR PROSITE: PS00251; TNF\_1; 1.

DR PROSITE: PS50049; TNF\_2; 1.

KW Cytokine; Cytotoxin; Glycoprotein.

FT NON\_TER 1 1

FT CARBOHYD 128 128

SQ SEQUENCE 150 AA; 16423 MW; FE5C4CC657658B48 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 150;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 194 GXYLYC 200  
DB 37 GXYLYC 43

RESULT 21

RA05\_ORYSA

ID RA05\_ORYSA STANDARD: PRT; 157 AA.

AC 001861;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-JUN-1994 (Rel. 29, Last annotation update)

DE Seed allergenic protein RA5 precursor.

GN RA5.

OS Oryza sativa (Rice).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Eriaristidae; Oryzae; Oryza.

OK NCBI\_TaxID=4530;

RN (1)

RP SEQUENCE FROM N.A.

RT TISSUE=Seed;

RX MEDLINE=93144699; PubMed=7678765;

RA Adachi T., Izumi H., Yamada T., Tanaka K., Takeuchi S.,

RA Nakamura R., Matsuda T.;

RT "Gene structure and expression of rice seed allergenic proteins  
RT belonging to the alpha-amylase/trypsin inhibitor family.";  
RT Plant Mol. Biol. 21:239-248(1993).  
CC -1- PTM: FIVE DISULFIDE BONDS ARE PRESENT (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE CEREAL TRYPsin/ALPHA-AMYLASE INHIBITOR  
CC FAMILY.

```

CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D11430; BAA01996.1;
CC PIR: S31078; S31078.
CC HSSP: P01085; 1HSS.
CC InterPro: IPR003612; AAI.
CC InterPro: IPR001768; try/amy1_inhbt.
CC Pfam: PF00234; try_alpha_amy1; 1.
CC PRINTS: PR00808; AMLASINBTR.
CC SMART: SM00499; AAI; 1.
CC PROSITE: PS00426; CEREAL_TRYP_AMYL_INH; 1.
CC Allergen: MultiGene family; Signal.
CC SIGNAL
CC CHAIN 1 26 SEED ALLERGENIC PROTEIN RAS.
CC SEQUENCE 157 AA; 17118 MW; C8A5495FBFB399E6 CRC64;
CC -----
OY 72 LLLAVS 78
Db 11 LLLAVS 17
-----
Query Match 2.58; Score 7; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
-----
RESULT 22
HLLP_ECOLI STANDARD; PRT; 161 AA.
ID HLLP_ECOLI
AC P11457;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Histone-like protein HLP-1 precursor (DNA-binding 17 kDa protein).
GN HLLP OR SKP OR OMPH OR B0178 OR Z0190 OR ECS0180.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 21-40.
RX MEDLINE=86329735; PubMed=2843433;
RA HOLCK A., Kleppe K.;
RT "Cloning and sequencing of the gene for the DNA-binding 17K protein
RT of Escherichia coli."
RL Gene 67:117-124(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.D.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN-K12 / W3110;
RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
RA Mizuno T., Makino K., Nakata A., Yura T., Samped G., Mizobuchi K.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of the
RT 4.0 - 5.0 min (189,987 - 281,416bp) region."
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [4]

```

```

RP SEQUENCE FROM N.A.
RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RA Davis K., Federpiel N., Hyman R., Kalman S., Komp C., Kurl O.,
RA Lashari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
RA Davis R.W.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RX STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7."
RL Nature 409:529-533(2001).
RN [6]
RP SEQUENCE FROM N.A.
RX STRAIN-O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shinagawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12."
RL DNA Res. 8:11-22(2001).
RN [7]
RP SEQUENCE OF 72-161 FROM N.A.
RX STRAIN-K12 / MG1655;
RX MEDLINE=91100302; PubMed=1987124;
RA Dicker I.B., Seetharam S.R.;
RT "Cloning and nucleotide sequence of the fira gene and the fira200(Ts)
RT allele from Escherichia coli."
RL J. Bacteriol. 173:334-344(1991).
RN [8]
RP SEQUENCE OF 21-32.
RX STRAIN-K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12."
RL Electrophoresis 18:1259-1313(1997).
RN [9]
RP SIMILARITY TO S. TYPHIMURUM OMPH.
RX MEDLINE=90201355; PubMed=2318304;
RA Hirvas L., Coleman J., Koski P., Vaara M.;
RT "Bacterial 'histone-like protein I' (HLP-I) is an outer membrane
RT constituent".
RL FEBS Lett. 262:123-126(1990).
CC -I- SUBUNIT: HOMOTETRAMER.
CC -I- SUBCELLULAR LOCATION: EITHER IN THE NUCLEOID (CHROMATIN) OR IN
CC THE OUTER MEMBRANE.
CC -I- SIMILARITY: BELONGS TO THE OMPH/HLLP FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M21118; AAA24630.1;
CC EMBL: AE000127; AAC73289.1;
CC EMBL: D83536; BAA77853.1;
CC EMBL: U70214; AAB08607.1;
CC EMBL: AE005194; AAG54460.1;
CC EMBL: AF002550; BAB33603.1;
CC EMBL: X54797; CAA38567.1;
CC EMBL: X75465; CAA53207.1;

```



DR PIR: J70304; DNEC17.  
 DR PIR: S13728; S13728.  
 DR Ecogene: EG10455; h1p.  
 KM DNA-binding; Outer membrane; Signal; Complete proteome.  
 FT SIGNAL 1 20  
 FT CHAIN 21 161 HISTONE-LIKE PROTEIN HLP-1.  
 SQ SEQUENCE 161 AA; 17688 MW; 2A965BBD83F3E675 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 34;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 GIGLALA 68  
 DB 9 GIGLALA 15

RESULT 23  
 YIM4\_CAEEL STANDARD; PRT; 170 AA.  
 AC P34378:  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Hypothetical protein D2007.4 in chromosome III.  
 GN D2007.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea;  
 OC Rhabditiidae; Pelodertinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Bristol N2;  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Almscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Gratton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,  
 RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,  
 RA Johnson L., Jones M., Kershaw J., Kirsten J., Lalister N.,  
 RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifkin L., Koopra A., Saunders D., Showkeen R.,  
 RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,  
 RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,  
 RA Waterston R., Watson A., Wellstock L., Wilkinson-Sproat J.,  
 RA Wohldman P.;  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 RT elegans.";  
 RL Nature 368:32-38(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: L16560; AAA27999.1; -  
 DR PIR: S44789; S44789.  
 DR WormPep: D2007.4; CE00129.  
 KM Hypothetical protein.  
 SQ SEQUENCE 170 AA; 19396 MW; 22301D7C65638135 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 170;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 221 VIALRCL 227  
 DB 94 VIALRCL 100

RESULT 24

LACB\_BUBBU STANDARD; PRT; 180 AA.  
 AC P02755; O62822;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Beta-lactoglobulin precursor (Beta-Lg).  
 GN LGB.  
 OS Bubalus bubalis (Domestic water buffalo).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bubalus.  
 OX NCBI\_TaxID=89462;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Mammary gland;  
 RX MEDLINE=99304500; PubMed=10376212;  
 RA Das P., Jain S., Nayak S., Apparao K.B.C., Toley S.M., Gary L.C.;  
 RT "Molecular cloning and sequence analysis of the cDNA encoding  
 RT beta-lactoglobulin in Bubalus bubalis.";  
 RL DNA Seq. 10:105-108(1999).  
 [2]  
 RN SEQUENCE OF 19-180.  
 RA Kolde H.-J., Liberatori J., Braunltzer G.;  
 RT "The amino acid sequence of the water buffalo beta-lactoglobulin.";  
 RL Milchwissenschaft 36:83-86(1981).

CC -1- FUNCTION: Primary component of whey, it binds retinol and is  
 CC probably involved in the transport of that molecule.  
 CC -1- SUBUNIT: UNDER PHYSIOLOGICAL CONDITIONS BETA-LACTOGLOBULIN EXISTS  
 CC AS AN EQUILIBRIUM MIXTURE OF MONOMERIC AND DIMERIC FORMS.  
 CC -1- SUBCELLULAR LOCATION: Secreted.  
 CC -1- TISSUE SPECIFICITY: Synthesized in mammary gland and secreted  
 CC in milk.

CC -1- MISCELLANEOUS: ALTERNATE DISULFIDE BONDS OCCUR IN EQUAL AMOUNTS.  
 CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AF005429; CA06532.1; -  
 DR PIR: A03219; LGBU.  
 DR HSSP: P02754; LB50.  
 DR InterPro: IPR002345; Lipocalin.  
 DR Pfam: PF00061; Lipocalin\_1.  
 DR PRINTS: PR00179; LIPOCALIN.  
 DR PROSITE: PS00213; LIPOCALIN; 1.  
 FT SIGNAL 1 18  
 FT CHAIN 19 180 BETA-LACTOGLOBULIN.  
 FT DISULFID 84 178  
 FT DISULFID 124 137  
 FT DISULFID 124 139  
 SQ SEQUENCE 180 AA; 20023 MW; 6836C97B2C2E33CF CRC64;

Query Match 2.5%; Score 7; DB 1; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGIALAC 69  
 DB 8 LGIALAC 14

RESULT 25  
 LACB\_CAPHI STANDARD; PRT; 180 AA.  
 ID LACB\_CAPHI  
 AC P02756;

DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Beta-lactoglobulin precursor (Beta-Lg).  
 GN LGB.  
 OS Capra hircus (Goat).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Caprinae; Capra.  
 OK NCBI\_TaxID=9925;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SP. aegagrus; TISSUE=Mammary gland;  
 RX MEDLINE=94042559; PubMed=8226387;  
 RA Folch J.M., Coll A., Sanchez A.;  
 RT "Cloning and sequencing of the cDNA encoding goat beta-lactoglobulin."  
 RL J. Anim. Sci. 71:2832-2832(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Kim J., Kim A., Kim J., Yu M.;  
 RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=95213451; PubMed=7699130;  
 RA Folch J., Coll A., Sanchez A.;  
 RT "Complete sequence of the caprine beta-lactoglobulin gene."  
 RL J. Dairy Sci. 77:3493-3497(1994).  
 RN [4]  
 RP SEQUENCE OF 19-180.  
 RX MEDLINE=80070611; PubMed=511095;  
 RA Preau G., Braunitzer G., Schrank B., Strangl A.;  
 RT "The amino acid sequence of goat beta-lactoglobulin."  
 RL Hoppe-Seyler's Z. Physiol. Chem. 360:1595-1604(1979).  
 CC -1 FUNCTION: Primary component of whey, it binds retinol and is probably involved in the transport of that molecule.  
 CC -1 SUBUNIT: UNDER PHYSIOLOGICAL CONDITIONS BETA-LACTOGLOBULIN EXISTS AS AN EQUILIBRIUM MIXTURE OF MONOMERIC AND DIMERIC FORMS.  
 CC -1 SUBCELLULAR LOCATION: Secreted.  
 CC -1 TISSUE SPECIFICITY: Synthesized in mammary gland and secreted in milk.  
 CC -1 MISCELLANEOUS: ALTERNATE DISULFIDE BONDS OCCUR IN EQUAL AMOUNTS.  
 CC -1 SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/announce/](http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch) or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X58471; CAA41385.1; -  
 DR EMBL: Z19569; CAA79623.1; -  
 DR EMBL: Z19570; CAA79624.1; -  
 DR EMBL: Z33881; CAA83946.1; -  
 DR PIR: A03320; LGGT.  
 DR PIR: S14507; S14507.  
 DR PIR: S42800; S42800.  
 DR HSP: P02754; IBOO.  
 DR InterPro: IPR002345; LIPOCALIN.  
 DR InterPro: IPR000566; LIPOCALIN\_CYTFAFP.  
 DR Pfam: PF00061; LIPOCALIN.1.  
 DR PRINTS: PR00179; LIPOCALIN.  
 DR PROSITE: PS00213; LIPOCALIN; 1.  
 DR MILK: Whey; Retinol-binding; Transport; LIPOCALIN; Signal.  
 FT CHAIN 1 180  
 FT SIGNAL 1 18  
 FT DISULFID 84 178 BETA-LACTOGLOBULIN.  
 FT DISULFID 124 137  
 FT DISULFID 124 139  
 SO SEQUENCE 180 AA; 19975 MW; C2449BB02A1A80F1 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 180;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 63 LGALAC 69  
 Db 8 LGALAC 14  
 RESULT 26  
 LACB-SHEEP  
 ID LACB-SHEEP STANDARD; PRT; 180 AA.  
 AC P02757;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-AUG-1988 (Rel. 08, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Beta-lactoglobulin 1/B, 2/A, and 3/C precursor.  
 OS Ovis aries (Sheep), and  
 OS Ovis orientalis musimon (Mouflon).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Caprinae; Ovis.  
 OK NCBI\_TaxID=9940, 9938;  
 RN [1]  
 RP SEQUENCE FROM N.A. (BLG 1 AND 2).  
 RC SPECIES=Sheep;  
 RX MEDLINE=88172489; PubMed=3351935;  
 RA Ali S., Clark A.J.;  
 RT "Characterization of the gene encoding ovine beta-lactoglobulin. Similarity to the genes for retinol binding protein and other secretory proteins."  
 RL J. Mol. Biol. 199:415-426(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A. (BLG 1).  
 RC SPECIES=Sheep;  
 RX MEDLINE=8704827; PubMed=3096387;  
 RA Gaye P., Hue-Delehante D., Mercier J.-C., Soulier S., Vilotte J.-L., Furet J.-P.;  
 RT "Ovine beta-lactoglobulin messenger RNA: nucleotide sequence and mRNA levels during functional differentiation of the mammary gland."  
 RL Biochimie 68:1097-1107(1986).  
 RN [3]  
 RP SEQUENCE FROM N.A. (BLG 1).  
 RC SPECIES=Sheep;  
 RX MEDLINE=89057492; PubMed=3194215;  
 RA Harris S., Ali S., Anderson S., Archibald A.L., Clark A.J.;  
 RT "Complete nucleotide sequence of the genomic ovine beta-lactoglobulin gene."  
 RL Nucleic Acids Res. 16:10379-10380(1988).  
 RN [4]  
 RP SEQUENCE FROM N.A. (BLG 1 AND 2).  
 RC SPECIES=Sheep;  
 RX MEDLINE=91007276; PubMed=1976573;  
 RA Ali S., McElenaghan M., Simons J.P., Clark A.J.;  
 RT "Characterisation of the alleles encoding ovine beta-lactoglobulins A and B."  
 RL Gene 91:201-207(1990).  
 RN [5]  
 RP SEQUENCE OF 19-180 (BLG 2).  
 RC SPECIES=Sheep;  
 RX MEDLINE=80219294; PubMed=6155855;  
 RA Preau G., Braunitzer G., Kolde H.-J.;  
 RT "Primary structure of ovine beta-lactoglobulin."  
 RL Arch. Int. Physiol. Biochim. 88:B45-B46(1980).  
 RN [6]  
 RP SEQUENCE OF 19-180 (BLG 3).  
 RC SPECIES=Sheep;  
 RX MEDLINE=89374823; PubMed=2775495;  
 RA Erhardt G., Godovac-Zimmermann J., Conti A.;  
 RT "Isolation and complete primary sequence of a new ovine wild-type beta-lactoglobulin C."  
 RL Biol. Chem. Hoppe-Seyler 370:757-762(1989).

```

RN [7]
RP SEQUENCE OF 19-180 (BIG B).
RC SPECIES=O. musimom; PubMed=3426802;
RX MEDLINE=88106996; PubMed=3426802;
RA Godovac-Zimmermann J., Conti A., Napolitano L.;
RT "The complete amino-acid sequence of dimeric beta-lactoglobulin from
  mouflon (Ovis ammon musimon) milk.";
RL Biol. Chem. Hoppe-Seyler 368:1313-1319(1987).
CC -1- FUNCTION: LACTOGLOBULIN IS THE PRIMARY COMPONENT OF WHEY, IT
  BINDS RETINOL AND IS PROBABLY INVOLVED IN THE TRANSPORT OF
  THAT MOLECULE.
CC -1- SUBUNIT: UNDER PHYSIOLOGICAL CONDITIONS BETA-LACTOGLOBULIN EXISTS
  AS AN EQUILIBRIUM MIXTURE OF MONOMERIC AND DIMERIC FORMS.
CC -1- MISCELLANEOUS: ALTERNATE DISULFIDE BONDS OCCUR IN EQUAL AMOUNTS.
CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X04520; CAA28204.1; -
DR EMBL: X12817; CAA31305.1; -
DR EMBL: X07004; CAA30059.1; ALT. SEQ.
DR EMBL: X07005; CAA30059.1; JOINED.
DR EMBL: X07006; CAA30059.1; JOINED.
DR EMBL: X07007; CAA30059.1; JOINED.
DR EMBL: X07008; CAA30059.1; JOINED.
DR EMBL: X07009; CAA30059.1; JOINED.
DR EMBL: M32236; AAA31510.1; -
DR EMBL: M32232; AAA31510.1; JOINED.
DR EMBL: M32233; AAA31510.1; JOINED.
DR EMBL: M32235; AAA31510.1; JOINED.
DR PIR: A03221; LGSH.
DR PIR: S00132; S00132.
DR PIR: A30011; A30011.
DR PIR: B30011; B30011.
DR PIR: A25136; A25136.
DR PIR: J00748; J00748.
DR PIR: J00749; J00749.
DR PIR: S02136; S02136.
DR PIR: S04955; S04955.
DR HSSP: P02754; 1BSO.
DR InterPro: IPR002345; Lipocalin.
DR Pfam: PF000566; Lipocalin_cytfabp.
DR PRINTS: PR00179; LIPOCALIN.
DR PROSITE: PS00213; LIPOCALIN.
KW Milk; Whey; Retinol-binding; Transport; Signal; Lipocalin.
FT SIGNAL 1 18
FT CHAIN 19 180
FT DISULFID 84 178
FT DISULFID 124 137
FT DISULFID 124 139
FT VARIANT 38 38
FT VARIANT 166 166
FT SEQUENCE 180 AA; 19921 MW; BABC82E89E757333 CRC64;
  Query Match 2.5%; Score 7; DB 1; Length 180;
  Best Local Similarity 100.0%; Pred. No. 38;
  Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 63 LGALAC 69
  |||||
  DB 8 LGALAC 14
  |||||
RESULT 27
RNF_E_SALTY STANDARD: PRT; 230 AA.
  ID RNF_E_SALTY

```

```

AC O8XEX9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Electron transport complex protein rnfE.
GN RNF_E OR STM1454 OR STY1668.
OS Salmonella typhimurium, and
  Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
  Salmonella.
OC NCBI_TaxID=602, 601;
OX [1]
RN SEQUENCE FROM N.A.
RP SPECIES=S. typhimurium; STRAIN=LT2 / SGSC1412 / ATCC 700720;
RC MEDLINE=21534948; PubMed=11677609;
RX McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
  Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D.,
  Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
  Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
  Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
  LT2.";
RL Nature 413:852-856(2001).
RN [2]
RN SEQUENCE FROM N.A.
RP SPECIES=S. typhi; STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
  Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
  Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
  Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
  Felwell T., Hamlin N., Haque A., Hien T.T., Holtroyd S., Jagels K.,
  Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
  Quail M., Rutherford K., Simmonds M., Skellon J., Stevens K.,
  Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
  enterica serovar Typhi CT18.";
RL Nature 413:848-852(2001).
CC -1- FUNCTION: May be part of a membrane complex involved in electron
  transport (By similarity).
CC -1- SUBUNIT: Composed of at least six subunits: rnfA, rnfB, rnfC,
  rnfD, rnfE and rnfG (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
  (Potential).
CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (see http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE008763; AAL20376.1; -
DR EMBL: AL627271; CAD01913.1; -
DR StyGene: SG27272; rnfE.
DR InterPro: IPR003667; Rnf_Mgr.
DR Pfam: PF02508; Rnf_Mgr.1.
KW Electron transport; Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 34 56
FT TRANSMEM 69 87
FT TRANSMEM 91 113
FT TRANSMEM 126 148
FT TRANSMEM 183 205
FT SEQUENCE 230 AA; 24318 MW; E198B4CEA13F249E CRC64;
  Query Match 2.5%; Score 7; DB 1; Length 230;
  Best Local Similarity 100.0%; Pred. No. 46;
  Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 60 ALGLCIA 66
  |||||

```

Db 38 ALGGLA 44

RESULT 28

RNFE\_ECO57 STANDARD; PRT; 231 AA.

AC P58344;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DE Electron transport complex protein rnfe.

GN RNFE OR Z2642 OR ECS2341.

OS Escherichia coli O157:H7.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

OX NCBI\_TaxID=83334;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-O157:H7 / EDL933 / ATCC 700927;

RX MEDLINE=21074935; PubMed=11206551;

RA Perna N.T., Plunkett G., Iii, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grodeck E.V., Davis N.W., Lam A., Dimantanta E.T., Potamouls K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;

RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7." Nature 409:529-533(2001).

RL [2]

RN SEQUENCE FROM N.A.

RC STRAIN-O157:H7 / RIMD 0509952;

RX MEDLINE=21156231; PubMed=11258796;

RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T., Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T., Kubara S., Shiba T., Hattori M., Shinagawa H.;

RT "Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12." DNA Res. 8:11-22(2001).

RL -1- FUNCTION: May be part of a membrane complex involved in electron transport (By similarity).

CC -1- SUBUNIT: Composed of at least six subunits; rnfa, rnfb, rnfc, rnfd, rnfe and rnfg (By similarity).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (potential).

CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AE005386; AAG56621.1; -

DR EMBL; AP002558; BAB35764.1; -

DR InterPro; IPR003667; Rnf\_Ngr.

DR Pfam; PF02508; Rnf\_Ngr; 1.

KW Electron transport; Transmembrane; Inner membrane; Complete proteome.

FT DOMAIN 1 38 PERIPLASMIC (POTENTIAL).

FT TRANSSEM 39 59 POTENTIAL.

FT DOMAIN 60 62 CYTOPLASMIC (POTENTIAL).

FT TRANSSEM 63 83 POTENTIAL.

FT DOMAIN 84 85 PERIPLASMIC (POTENTIAL).

FT TRANSSEM 86 106 POTENTIAL.

FT DOMAIN 107 124 CYTOPLASMIC (POTENTIAL).

FT TRANSSEM 125 145 POTENTIAL.

FT DOMAIN 146 181 PERIPLASMIC (POTENTIAL).

FT TRANSSEM 182 202 POTENTIAL.

FT DOMAIN 203 231 CYTOPLASMIC (POTENTIAL).

SQ SEQUENCE 231 AA; 24489 MW; DAA2CA2D292604C3 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 231;

Best Local Similarity 100.0%; Pred. No. 47;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGGLA 66

Db 38 ALGGLA 44

RESULT 29

RNFE\_ECOLI STANDARD; PRT; 231 AA.

AC P71719;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DE Electron transport complex protein rnfe.

GN RNFE OR B1632.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

OX NCBI\_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;

RX MEDLINE=97426617; PubMed=9278503;

RA Blattner F.R., Plunkett G., Iii, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Mau B., Shao Y.;

RT "The complete genome sequence of Escherichia coli K-12." Science 277:1453-1474(1997).

RL [2]

RN SEQUENCE FROM N.A.

RC STRAIN-K12;

RX MEDLINE=97251357; PubMed=9097039;

RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K., Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M., Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Oshima T., Saito N., Sampei G., Seki Y., Sivasubraman S., Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C., Yamamoto Y., Horiuchi T.;

RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 28.0-40.1 min region on the linkage map." DNA Res. 3:363-377(1996).

RL [3]

RN TOPOLOGY.

RC STRAIN-K12 / JM109;

RX MEDLINE=99342054; PubMed=10411911;

RA Seeat A., Johansson M., Wallin E., von Heijne G.;

RT "Divergent evolution of membrane protein topology: the Escherichia coli rnfa and rnfe homologues." Proc. Natl. Acad. Sci. U.S.A. 96:8540-8544(1999).

RL -1- FUNCTION: May be part of a membrane complex involved in electron transport (By similarity).

CC -1- SUBUNIT: Composed of at least six subunits; rnfa, rnfb, rnfc, rnfd, rnfe and rnfg (By similarity).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (potential).

CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AE000258; AAC74704.1; -

DR EMBL; D90806; BAA15386.1; -

DR EMBL; D90807; BAA15393.1; -

DR EMBL: D90808; BAA15416.1; -  
 DR ECGene: EG13938; rnfE.  
 DR InterPro: IPR003667; Rnf\_Ngr.  
 DR Pfam: PF02508; Rnf\_Ngr; 1.  
 KW Electron transport; Transmembrane; Inner membrane; Complete proteome.  
 FT DOMAIN 1 38 PERIPLASMIC (POTENTIAL).  
 FT TRANSMEM 39 59 POTENTIAL.  
 FT DOMAIN 60 62 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 63 83 POTENTIAL.  
 FT DOMAIN 84 85 PERIPLASMIC (POTENTIAL).  
 FT TRANSMEM 86 106 POTENTIAL.  
 FT DOMAIN 107 124 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 125 145 POTENTIAL.  
 FT DOMAIN 146 181 PERIPLASMIC (POTENTIAL).  
 FT TRANSMEM 182 202 POTENTIAL.  
 FT DOMAIN 203 231 CYTOPLASMIC (POTENTIAL).  
 SQ SEQUENCE 231 AA; 24459 MW; CFA37A2D92604C3 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 231;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGGLA 66  
 |||||  
 DB 38 ALGGLA 44

## RESULT 30

RNC\_CONXBU STANDARD; PRT; 233 AA.

AC P51837;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ribonuclease III (EC 3.1.26.3) (Rnase III).  
 GN Rnc.  
 OS Coccidia burnetii.  
 OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;  
 CC Coccidia group; Coccidia.  
 OX NCBI\_TaxID=777;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBM1C7;  
 RX MEDLINE=95131751; PubMed=7830573;  
 RA Zuber M., Hoover T.A., Powell B.S., Court D.L.;  
 RL Mol. Microbiol. 14:291-300(1994).  
 CC -1- FUNCTION: DIGESTS DOUBLE-STRANDED RNA. INVOLVED IN THE PROCESSING  
 OF RIBOSOMAL RNA PRECURSORS AND OF SOME MRNAs (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 phosphomonoester.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: CONTAINS 1 DBM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 RNASE III DOMAIN.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: L27436; AAA69690.1; -  
 DR InterPro: IPR001159; DS\_RBD.  
 DR Pfam: PF00035; dsrm; 1.  
 DR Pfam: PF00636; Ribonuclease\_3; 1.  
 DR SMART: SM00358; DSRM; 1.  
 DR SMART: SM00535; RIBOC; 1.  
 DR PROSITE: PS50137; DS\_RBD; 1.  
 DR PROSITE: PS00517; RNASE\_3\_1; 1.  
 DR PROSITE: PS50142; RNASE\_3\_2; 1.

KW Hydrolase; Nuclease; Endonuclease; RNA-binding.  
 FT DOMAIN 4 126 RNASE III.  
 FT TRANSMEM 204 220 DBM.  
 SQ SEQUENCE 233 AA; 26229 MW; 1A11C84AD96ED2FF CRC64;

Query Match 2.5%; Score 7; DB 1; Length 233;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 ARRLPL 16  
 |||||  
 DB 164 ARRLPL 170

## RESULT 31

RNFE\_YERPE STANDARD; PRT; 233 AA.

AC Q8ZED4;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Electron transport complex protein rnfE.  
 GN RNFE OR YP02240.  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Yersinia.  
 OX NCBI\_TaxID=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CO-92 / Biovar Orientalis;  
 RX MEDLINE=21470413; PubMed=11586360;  
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,  
 Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,  
 Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
 Feltham S., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,  
 Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
 Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;  
 RA "Genome sequence of Yersinia pestis, the causative agent of plague."  
 RL Nature 413:523-527(2001).  
 CC -1- FUNCTION: May be part of a membrane complex involved in electron  
 transport (By similarity).  
 CC -1- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC,  
 rnfD, rnfE and rnfG (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AJ14151; CAC91046.1; -  
 DR InterPro: IPR003667; Rnf\_Ngr.  
 DR Pfam: PF02508; Rnf\_Ngr; 1.  
 KW Electron transport; Transmembrane; Inner membrane; Complete proteome.  
 FT TRANSMEM 34 56 POTENTIAL.  
 FT TRANSMEM 69 87 POTENTIAL.  
 FT TRANSMEM 91 113 POTENTIAL.  
 FT TRANSMEM 126 148 POTENTIAL.  
 FT TRANSMEM 184 206 POTENTIAL.  
 SQ SEQUENCE 233 AA; 24587 MW; 491E18F335E8CB90 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 233;  
 Best Local Similarity 100.0%; Pred. No. 47;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGGLA 66  
 |||||

```

Db      38 ALGIGLA 44

RESULT 32
RNF_E_HAEIN STANDARD; PRT: 235 AA.
AC Q57020; P96346;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Electron transport complex protein rnfE.
OS RNF_E OR H11688.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_Taxid=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Sherry R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Usterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fritchman J.L., Geoghagen N.S.M.,
RA Grehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd."
RL Science 269:496-512(1995).
CC -1- FUNCTION: May be part of a membrane complex involved in electron
CC transport (By similarity).
CC -1- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC,
CC rnfD, rnfE and rnfG (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32841; AAC23334.1; -
DR TIGR: H11688; -
DR InterPro: IPR003667; Rnf_Nqr.
DR Pfam: PF02508; Rnf_Nqr; 1.
KW Electron transport; Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 63 83 POTENTIAL.
FT TRANSMEM 93 113 POTENTIAL.
FT TRANSMEM 117 137 POTENTIAL.
FT TRANSMEM 152 172 POTENTIAL.
FT TRANSMEM 206 226 POTENTIAL.
SQ SEQUENCE 235 AA; 25845 MW; C054FE596647837A CRC64;

Query Match 2.58; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 60 ALGIGLA 66
Db 62 ALGIGLA 68

RESULT 33
TN14_MOUSE STANDARD; PRT: 239 AA.
AC Q90YH9;

DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 14.
OS TNFSF14 OR LIGHT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=10700230;
RA Tamada K., Shimozaki K., Chapoval A.I., Zhu G., Sica G., Flies D.,
RA Boone T., Hsu H., Fu Y.-X., Nagata S., Ni J., Chen L.;
RT "Modulation of T-cell-mediated immunity in tumor and graft-versus-host
RT disease models through the LIGHT co-stimulatory pathway.";
RL Nat. Med. 6:283-289(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Fetal liver;
RA Misawa K., Nosaka T., Kitamura T., Kojima T.;
RT "Murine LIGHT, a homologue of human LIGHT which is a member of TNF
RT family.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Lymphoma;
RA Force W.R., Todd P.K., Mikayama T.;
RT "Mouse LIGHT: molecular genetics, ligand binding and expression.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cytokine that binds to TNFRSF3/LTBR. Binding to the
CC decoy receptor TNFRSF6B modulates its effects. Activates NFkB
CC and stimulates the proliferation of T cells.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
CC similarity).
CC -1- PTM: The soluble form derives from the membrane form by
CC proteolytic processing.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF123385; AAF6453.1; -
DR EMBL: AB029155; BAA8859.1; -
DR EMBL: AF227533; AAF36722.1; -
DR HSSP: P01375; 4TSV.
DR MGDI: MG11355317; Tnfsl4.
DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR004478; TNF_family.
DR Pfam: PF00229; TNF; 1.
DR PRINTS: PR01234; TNECROSISFCT.
DR PRODOM: PD002012; TNF_abc; 1.
DR SMART: SM00207; TNF; 1.
DR PROSITE: PS00251; TNF; 1; FALSE-NEG.
DR PROSITE: PS50049; TNF_2; 1.
KW Cytokine; Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor.
FT CHAIN 1 239 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
FT CHAIN 782 239 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
FT CHAIN 1 239 MEMBER 14, MEMBRANE FORM.
FT TRANSMEM 38 58 MEMBER 14, SOLUBLE FORM.
FT TRANSMEM 59 239 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 81 82 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT SITE 152 187 (POTENTIAL).
FT DISULFID 100 100 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 100 100 CLEAVAGE (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 191 191 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 239 AA; 26338 MW; 217874AC71ADBEB3 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 239;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 121 NRLVPR 127  
Db 221 NRLVPR 227

RESULT 34  
RNF\_E\_PSEAE STANDARD; PRT; 240 AA.

AC Q9HIB5;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Electron transport complex protein rnfE.  
GN RNF OR PA3494.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 15692 / PA01;  
RA MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen.";  
RL Nature 406:959-964(2000).

CC -1- FUNCTION: May be part of a membrane complex involved in electron  
transport (By similarity).  
CC -1- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC,  
CC rnfD, rnfE and rnfG (By similarity).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Potential).  
CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AE004770; AAC06882.1; -  
DR InterPro: IPR003667; Rnf\_Ngr.  
DR Pfam: PF02508; Rnf\_Ngr; 1.  
DR Electron transport; Transmembrane; Inner membrane; Complete proteome.  
FT TRANSMEM 41 61 POTENTIAL.  
FT TRANSMEM 71 91 POTENTIAL.  
FT TRANSMEM 95 115 POTENTIAL.  
FT TRANSMEM 130 150 POTENTIAL.  
FT TRANSMEM 184 204 POTENTIAL.  
SQ SEQUENCE 240 AA; 25794 MW; 3D90687ED462D8B2 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 240;  
Best Local Similarity 100.0%; Pred. No. 48;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLA 66  
Db 40 ALGLA 46

RESULT 35

ID MOEB\_HAEIN STANDARD; PRT; 243 AA.  
AC P45211;

DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Molybdoxin biosynthesis protein moeb.  
GN MOEB OR CHLN OR H1449.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.

OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Rd / KW20 / ATCC 51907;  
RA MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
RL Rd.";  
RL Science 269:496-512(1995).

CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A MOLYBDENUM COPROFACOR  
(MOLYBDOPROTEIN). NECESSARY FOR MOLYBDENZYMES. PLAYS A ROLE IN THE  
CC ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOPROTEIN CONVERTING  
CC FACTOR (MOAF) (BY SIMILARITY).  
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.  
CC -1- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: U32823; AAC3099.1; -  
DR TIGR: H11449; -  
DR InterPro: IPR00205; NAD\_binding.  
DR InterPro: IPR000594; THIF\_domain.  
DR Pfam: PF00899; Thif; 1.  
DR Molybdenum cofactor biosynthesis; Complete proteome.  
SQ SEQUENCE 243 AA; 26996 MW; 218A3382A975BDBD CRC64;

Query Match 2.5%; Score 7; DB 1; Length 243;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 179 LRYNROI 185  
Db 10 LRYNROI 16

RESULT 36  
TNFC\_HUMAN STANDARD; PRT; 244 AA.

AC Q0643; P78370; Q99761;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Lymphotoxin-beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor  
DE necrosis factor ligand superfamily member 3).  
GN LTB OR TNFSF3 OR TNFC.  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 [1]  
 RN SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.  
 RC TISSUE=T-cell;  
 RX MEDLINE=93208881; PubMed=7916655;  
 RA Browning J.L., Ngam-Ek A., Lawton P., Demarinis J., Tizard R.,  
 RA Chow E.P., Hession C., O'Brien-Greco B., Foley S.F., Ware C.F.,  
 RT "Lymphotoxin beta, a novel member of the TNF family that forms a  
 RT heteromeric complex with lymphotoxin on the cell surface."  
 RL Cell 72:847-856(1993).  
 [2]  
 RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RX MEDLINE=97445965; PubMed=9299492;  
 RA Marzocha K., Renard N., Charlot C., Blenvenu J., Collier B.,  
 RA Salles G.;  
 RT "Identification of two lymphotoxin beta isoforms expressed in human  
 RT lymphoid cell lines and non-Hodgkin's lymphomas."  
 RL Biochem. Biophys. Res. Commun. 238:273-276(1997).  
 [3]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Neville M.J., Milner C.M., Campbell R.D.;  
 RT "A new member of the immunoglobulin superfamily and a V-ATPase G  
 RT subunit are amongst the predicted products of novel genes close to the  
 RT TNF locus in the human MHC."  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 [4]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Rowen L., Madan A., Olin S., Shaffer T., James R., Ratcliffe A.,  
 RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dots M., Young J.,  
 RA Laszky S., Hood L.;  
 RT "Sequence of the human major histocompatibility complex class III  
 RT region."  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 [5]  
 RN SEQUENCE FROM N.A. (ISOFORM 1).  
 RA Shihina S., Tamura G., Oka A., Inoko H.;  
 RT "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region."  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 [6]  
 RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS GLD-70 AND  
 RP PRO-111.  
 RA Rieder M.J., Amel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,  
 RA Poel C.L., Toth E.J., Yi Q., Nickerson D.A.;  
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Cytokine that binds to LTRB/TNFRSF. May play a specific  
 CC role in immune response regulation. Provides the membrane anchor  
 CC for the attachment of the heterotrimeric complex to the cell  
 CC surface. Isoform 2 is probably non-functional.  
 CC -1- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits or  
 CC (less prevalent) one LTB and two LTA subunits.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Potential).  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2, are  
 CC produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: SPLEEN AND THYMUS.  
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: L11016; AAA99888.1; -  
 DR EMBL: 089922; AAC51769.1; -  
 DR EMBL: 079029; AAB37342.1; -  
 DR EMBL: L11015; AAA36191.1; -  
 DR EMBL: Y14768; CA75069.1; -  
 DR EMBL: AF129756; MAD18089.1; -  
 DR EMBL: AF000505; BAB63395.1; -

DR EMBL: AY070219; AAL49954.1; -  
 DR EMBL: AY070219; AAL49955.1; -  
 DR PIR: A46066; A46066.  
 DR HSSP: P01374; 1TNR.  
 DR Genew: HGNC:6711; LTB.  
 DR MIM: 600978; -  
 DR InterPro: IPR003636; TNF\_abc.  
 DR InterPro: IPR000478; TNF\_family.  
 DR Pfam: PF00229; TNF\_1.  
 DR PRINTS: PR01234; TNECROSISFCT.  
 DR ProDom: PD002012; TNF\_abc; 1.  
 DR SMART: SM00207; TNF; 1.  
 DR PROSITE: PS00251; TNF\_1; 1.  
 DR PROSITE: PS50049; TNF\_2; 1.  
 KW Cytokine; Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor;  
 KW Alternative splicing; Polymorphism.  
 FT DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 19 48 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT DOMAIN 49 222 EXTRACELLULAR (POTENTIAL).  
 FT CARBOHYD 222 222 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT VARSPPLIC 53 77 GLYTERADGAAGGGLGKPLPEE -> GLGPRSGQRSSO  
 FT VARSPPLIC 78 244 MISSING (IN ISOFORM 2).  
 FT VARIANT 70 70 G -> E.  
 FT VARIANT 111 111 /FTID-VAR\_013025.  
 FT FT A -> P.  
 FT FT /FTID-VAR\_013026.  
 FT FT DPGAAGGGL -> GLSAGSGGRT (IN REF. 2;  
 FT FT AAB37342)  
 SQ SEQUENCE 244 AA; 25390 MM; F4156459830ED4C CRC64;  
 Query Match 2.5%; Score 7; DB 1; Length 244;  
 Best Local Similarity 100.0%; Pred. No. 49;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 194 GLYLYYC 200  
 DB 131 GLYLYYC 137  
 RESULT 37  
 MOEB\_ECOLI  
 ID MOEB\_ECOLI STANDARD; PRT; 249 AA.  
 AC P12282;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Molydopterin biosynthesis protein moeb.  
 GN MOEB OR CHLN OR B0826.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OC NCBI\_TaxID=562;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=88314906; PubMed=3045084;  
 RA Nohno T., Kasai Y., Saito T.;  
 RT "Cloning and sequencing of the Escherichia coli chln operon involved  
 RT in molydopterin biosynthesis."  
 RL J. Bacteriol. 170:4097-4102(1988).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12."  
 RL Science 277:123-124(1997).  
 [3]  
 RN SEQUENCE FROM N.A.



```

RC STRAIN-K12;
RX OSHIMA-97061202; PubMed-8905232;
RA Oshima T., Alta H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horuchi T.,
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYDOP COFACTOR
CC (MOLYBDOPTERIN), NECESSARY FOR MOLYBDENZYMES. PLAYS A ROLE IN THE
CC ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOTERIN CONVERTING
CC FACTOR (MOAD).
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: M2151; AAA23580.1; -
DR EMBL: AE000185; AAC73913.1; -
DR EMBL: D90720; BAA35514.1; -
DR EMBL: D90721; BAA35521.1; -
DR PIR: B32352; B32352.
DR Ecogene: EG10154; MOEB.
DR InterPro: IPR000205; NAD_binding.
DR InterPro: IPR000594; THIF_domain.
DR Pfam: PF00899; Thif; 1.
SQ Molybdenum cofactor biosynthesis; Complete proteome.
SQ SEQUENCE 249 AA; 26719 MW; 12C77082B3F39D7D CRC64;

Query Match 2.5%; Score 7; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 LRYNRQI 185
DB 10 LRYNRQI 16

RESULT 38
MOEB_SALTY
ID MOEB_SALTY STANDARD; PRT; 249 AA.
AC 056057;
DR 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Molybdopterin biosynthesis protein moeb.
GN MOEB OR STM0845.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-LT2;
RA Wong K.K., Kwan H.S.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE-21534948; PubMed-11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Milvanev E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,

```

```

RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYDOP COFACTOR
CC (MOLYBDOPTERIN), NECESSARY FOR MOLYBDENZYMES. PLAYS A ROLE IN THE
CC ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOTERIN CONVERTING
CC FACTOR (MOAD).
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U53178; AAA96530.1; -
DR EMBL: AE008735; AAL19781.1; -
DR StyGene; SG1065; MOEB.
DR InterPro: IPR000205; NAD_binding.
DR InterPro: IPR000594; THIF_domain.
DR Pfam: PF00899; Thif; 1.
SQ Molybdenum cofactor biosynthesis; Complete proteome.
SQ SEQUENCE 249 AA; 26903 MW; 0F0050831D537AD2 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 LRYNRQI 185
DB 10 LRYNRQI 16

RESULT 39
KDPX_SERMA
ID KDPX_SERMA STANDARD; PRT; 257 AA.
AC Q54435;
DR 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Lipopolysaccharide core biosynthesis glycosyl transferase kdtX
DE (Ec 2.---.-).
GN KDPX.
OS Serratia marcescens.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Serratia.
OX NCBI_TaxID=615;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-N28B;
RX MEDLINE-96422003; PubMed-8824620;
RA Guasch J.F., Pique N., Climent N., Ferrer S., Merino S., Rubires X.,
RA Tomas J.M., Regue M.;
RT Cloning and characterization of two Serratia marcescens genes
RT involved in core lipopolysaccharide biosynthesis";
RL J. Bacteriol. 178:5741-5747(1996).
CC -1- PATHWAY: Lipopolysaccharide core biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2. WAAE/KDTX
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U52844; AAC44433.1; -  
DR InterPro: IPR001173; Glycosyltransferase.  
DR Pfam: PF00535; Glycosyltransferase; Transferrase; Glycosyltransferase.  
KW Lipopolysaccharide biosynthesis; Transferrase; Glycosyltransferase.  
SQ SEQUENCE 257 AA; 29233 MW; D40DB57E002FE90 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 257;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 71 GLLAUV 77  
|||||  
Db 229 GLLAUV 235

RESULT 40  
TNFC\_MARMO STANDARD; PRT; 310 AA.  
ID O3JML0: O3JML1;  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE Lymphotoxin-beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor  
DE necrosis factor ligand superfamily member 3).  
GN LTB OR TNFSF3 OR TNFC.  
OS Marmota monax (Woodchuck).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae;  
OC Marmota.  
OX NCBI\_TaxID=9995;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-20184748; PubMed-10721723;  
RA L.D.H., Havelle E.A., Brown C.L., Cullen J.M.;  
RT "Woodchuck lymphotoxin-alpha, -beta and tumor necrosis factor genes:  
RT structure, characterization and biological activity";  
RL Gene 242:295-305(2000).  
CC -1- FUNCTION: Cytokine that binds to LTB $\beta$ /TNFSF3. May play a specific  
CC role in immune response regulation. Provides the membrane anchor  
CC for the attachment of the heterotrimeric complex to the cell  
CC surface.  
CC -1- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits or  
CC (less prevalent) two LTA and one LTB subunits (By similarity).  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Potential).  
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF096268; AAF34866.1; -  
DR EMBL: AF095587; AAF34865.1; -  
DR HSSP: P01374; 1NR.  
DR InterPro: IPR003636; TNF-abc.  
DR InterPro: IPR004078; TNF-family.  
DR Pfam: PF00229; TNF; 1.  
DR PRINTS: PR01234; TNECROSISFCT.  
DR ProDom: PD002012; TNF-abc; 1.  
DR SMART: SM00207; TNF; 1.  
DR PROSITE: PSS0049; TNF\_2; 1.  
KW Cytokine; Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor.  
FT DOMAIN 1 27 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 28 48 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT DOMAIN 49 310 EXTRACELLULAR (POTENTIAL).

FT CARBOHYD 272 272 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CONFLICT 280 280 D -> H (IN REF. 1; AAF34865).  
SQ SEQUENCE 310 AA; 32644 MW; 73B354EFC8B3B3BE CRC64;

Query Match 2.5%; Score 7; DB 1; Length 310;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 194 GLYLVC 200  
|||||  
Db 181 GLYLVC 187

RESULT 41  
ISPH\_XYLFA STANDARD; PRT; 316 AA.  
ID ISPH\_XYLFA  
AC O9PAS9;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE ISPH protein.  
GN ISPH OR LTB OR XP2416.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OC Xylella.  
OX NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SRRAIN-9a5c;  
RX MEDLINE-20365717; PubMed-10910347;  
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvaranga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
RA Bueno M.R.P., Camargo A.I., Camargo L.E.A., Carro D.M., Carrier H.,  
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Fetto J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Hohnel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
RA Krieger J.E., Kuramae E.E., Laigret F., Lambals M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
RA Marques M.V., Martins A.L., Martins E.A.L., Matsushima A.Y.,  
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
RA Nani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsunako M.H.,  
RA Valada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
RA Zago M.A., Zatz M., Meidanis J., Seubert J.C.;  
RT "The genome sequence of the plant pathogen Xylella fastidiosa";  
RL Nature 406:151-159(2000).  
CC -1- FUNCTION: Converts 1-hydroxy-2-methyl-2-(E)-butenyl 4-diphosphate  
CC into isopentenyl diphosphate (IPP) and dimethylallyl diphosphate  
CC (DMAPP) (By similarity).  
CC -1- PATHWAY: Nonmevalonate terpenoid biosynthesis pathway; seventh  
CC (last) step.  
CC -1- SIMILARITY: BELONGS TO THE ISPH FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE004050; AAF85215.1; -  
 DR InterPro: IPR003451; LytB.  
 DR Pfam: PF02401; LytB; 1.  
 DR TIGRFAMs: TIGR00216; LytB; 1.  
 DR Isoprene biosynthesis; Complete proteome.  
 SQ SEQUENCE 316 AA; 34704 MW; 1A2E80B9A98D34A CRC64;

Query Match 2.5%; Score 7; DB 1; Length 316;  
 Best Local Similarity 100.0%; Pred. No. 61;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 217 LVDEYLA 223  
 |||||  
 Db 276 LVDEYLA 282

RESULT 42  
 ODPB\_BACSU STANDARD; PRT; 324 AA.

AC P21882;  
 DT 01-MAY-1991 (Rel. 18, Created)  
 DT 01-MAY-1991 (Rel. 18, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Pyruvate dehydrogenase E1 component, beta subunit (EC 1.2.4.1) (S  
 complex, 36 kDa subunit).  
 GN PDHB OR ACEB.  
 OS Bacillus subtilis.  
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=90368558; PubMed=1697575;  
 RA Hemila H., Palva A., Paulin L., Arvidsson S., Palva I.;  
 RT "Secretory S. complex of Bacillus subtilis: sequence analysis and  
 RT J. Bacteriol. 172:5052-5063(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=97124187; PubMed=8969500;  
 RA Winters P., Caldwell R., Enfield L., Ferrari E.;  
 RT "The amp<sup>s</sup>-np<sup>re</sup> (124 degrees-127 degrees) region of the Bacillus  
 RT subtilis 168 chromosome: sequencing of a 27 kb segment and  
 RT identification of several genes in the area.";  
 RL Microbiology 142:3033-3037(1996).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Caldwell R.M., Ferrari E.;  
 RT "Sequence analysis of the mobA-amps region of the Bacillus subtilis  
 RT chromosome.";  
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall  
 CC conversion of pyruvate to acetyl-CoA and CO(2). It contains  
 CC multiple copies of three enzymatic components: pyruvate  
 CC dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and  
 CC lipoamide dehydrogenase (E3).  
 CC -1- FUNCTION: THE B. SUBTILIS PDH COMPLEX POSSESSES ALSO BRANCHED-CHAIN  
 CC 2-OXOACID DEHYDROGENASE (BCDH) ACTIVITY.  
 CC -1- CATALYTIC ACTIVITY: Pyruvate + lipoamide = S-  
 CC acetylhydrolipoamide + CO(2).  
 CC -1- COFACTOR: THIAMINE PYROPHOSPHATE.  
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA CHAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M57435; AAA62682.1; -  
 DR EMBL: AF012285; AAC24933.1; -  
 DR EMBL: Z69111; CAB1332.1; -  
 DR PIR: C36718; C36718.  
 DR HSSP: P09061; I0S0.  
 DR Subtilist; BG10208; pdhb.  
 DR InterPro: IPR000360; Transketolase.  
 DR Pfam: PF02780; transket-pyr; 1.  
 DR Glycolysis; Oxidoreductase; Flavoprotein; Thiamine pyrophosphate;  
 KW Complete proteome.  
 FT INIT\_MET 0  
 SQ SEQUENCE 324 AA; 35343 MW; D2A7C9B32DED0F0D CRC64;

Query Match 2.5%; Score 7; DB 1; Length 324;  
 Best Local Similarity 100.0%; Pred. No. 62;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LALGIGL 65  
 |||||  
 Db 65 LALGIGL 71

RESULT 43  
 SRA6\_CAEEL STANDARD; PRT; 329 AA.

AC Q09208;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Serpentine receptor class alpha 6 (Sra-6 protein).  
 GN SRA-6 OR AH6.10.  
 OS Caenorhabditis elegans.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 CC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol NZ;  
 RA Jassal B.;  
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE C. ELEGANS RECEPTOR-LIKE PROTEIN SRA  
 CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z48009; CAA88083.1; -  
 DR WormPep; AH6.10; CE01451.  
 DR InterPro: IPR000344; Sra\_chemo\_Ce.  
 DR Pfam: PF02117; Sra; 1.  
 DR PRINTS: PR00697; TMPTROTEINSRA.  
 KW Transmembrane; Multigene family.  
 FT TRANSMEM 26 46  
 FT TRANSMEM 104 124  
 FT TRANSMEM 143 163  
 FT TRANSMEM 187 207  
 FT TRANSMEM 238 258  
 FT TRANSMEM 273 293  
 SQ SEQUENCE 329 AA; 37951 MW; 7C0963ADA53A29F3 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 329;  
 Best Local Similarity 100.0%; Pred. No. 63;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 134 RTRRAR 140  
 |||||

Db 297 RKRARR 303

RESULT 44

ING2\_RAT ID IAG2\_RAT STANDARD; PRT; 335 AA.

AC 035777;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Implantation-associated protein.  
GN IAG2.

OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_Taxid:10116;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN-Wistar; TISSUE-Uterus;  
RA Barkai U., Amit A., Leising J.B., Kraicer P.F., Kidron T.;  
RT "Implantation-associated uterine protein."  
RU Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- SIMILARITY: BELONGS TO THE OST3 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AF008554; AAB63284.2;  
DR InterPro: IPR000063; ThioRed.

KW Transmembrane.

FT TRANSMEM 6 26 POTENTIAL.

FT TRANSMEM 185 205 POTENTIAL.

FT TRANSMEM 210 230 POTENTIAL.

FT TRANSMEM 271 291 POTENTIAL.

FT TRANSMEM 301 321 POTENTIAL.

SQ SEQUENCE 335 AA; 37992 MW; AFB9DB35F1A06091 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 335;  
Best Local Similarity 100.0%; Pred. No. 64;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLAV 76  
|||||

Db 192 LGLLAV 198

RESULT 45

LEU3\_THEAO

ID LEU3\_THEAO STANDARD; PRT; 344 AA.

AC P24098;  
DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE 3-isopropylmalate dehydrogenase (EC 1.1.1.85) (Beta-IPM dehydrogenase)

DE (IMDH) (3-IPM-DH).  
GN LEUB.

OS Thermus aquaticus.  
OC Bacteria; Thermus; Deinococcus group; Deinococci; Thermales;  
OC Thermaceae; Thermus.

OX NCBI\_Taxid:271;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN-YT1;  
RX MEDLINE-92041736; PubMed-1939005;

RA Kirino H., Oshima T.;  
RT "Molecular cloning and nucleotide sequence of 3-isopropylmalate  
dehydrogenase gene (leub) from an extreme thermophile, Thermus

RT aquaticus YT-1.";  
RL J. Biochem. 109:852-857 (1991).

CC -1- CATALYTIC ACTIVITY: 3-carboxy-2-hydroxy-4-methylpentanoate +  
CC NAD(+) -> 3-carboxy-4-methyl-2-oxopentanoate + NADH.

CC -1- PATHWAY: Leucine biosynthesis; third step.

CC -1- SUBUNIT: HOMODIMER.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO THE ISOCITRATE AND ISOPROPYLMALATE  
CC DEHYDROGENASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: D10700; BAA01542.1;  
DR PIR: JX0173; JX0173.

DR HSSP: P00351; 1XNA.

DR InterPro: IPR001804; Isoch.

DR InterPro: IPR004429; Leub.

DR Pfam: PF00180; Isoch; 1.

DR TIGRFAMs: TIGR00169; leub; 1.

DR PROSITE: PS00470; IDH\_IMDH; 1.

KW Oxidoreductase; Leucine biosynthesis; NAD.

SQ SEQUENCE 344 AA; 36949 MW; 4F31A9444E826408 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 344;  
Best Local Similarity 100.0%; Pred. No. 66;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 248 SGLALR 254  
|||||

Db 88 SGLALR 94

Search completed: May 8, 2003, 06:56:24  
Job time : 40 secs

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 04:23:48 ; Search time 61 Seconds

(without alignments)  
620.380 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 284  
Sequence: 1 MSLLDFEISARLPLPRSLG.....PMAHLKAPFLTYGFLQVH 284

Scoring table:

Gapop 60.0, Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size: 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database:

A\_Geneseq\_101002:\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	284	100.0	284	19	AAW47525
2	249	87.7	249	20	AAW47525
3	249	87.7	249	21	AAW47525
4	249	87.7	249	23	AAW47525
5	249	87.7	249	23	AAW47525
6	241	84.9	249	19	AAW29745
7	241	84.9	249	22	AAW29745
8	207	72.9	273	22	AAW03499
9	146	51.4	146	22	AAW0895
10	143	50.4	189	19	AAW29746

11	143	50.4	189	22	AAW0892	Human UTAflap TREP
12	107	37.7	208	20	AAW3590	Human TMR13 protei
13	46	16.2	211	20	AAW3591	Mouse TMR13 protei
14	32	11.3	225	19	AAW47524	Mus musculus tumou
15	32	11.3	225	21	AAW07527	Amino acid sequenc
16	9	3.2	58	21	AAW01265	Human secreted pro
17	9	3.2	748	20	AAW14906	Human secreted pro
18	9	3.2	749	23	ABW73512	Extended amino aci
19	9	3.2	749	23	ABW73512	M vaccae GYS-9 pro
20	8	2.8	55	22	AAW1863	Propionibacterium
21	8	2.8	55	22	ABW42637	Peptide #10143 enc
22	8	2.8	55	22	ABW5983	Protein #7982 enco
23	8	2.8	55	22	AAW3528	Human brain expres
24	8	2.8	55	22	AAW76342	Human bone marrow
25	8	2.8	55	22	AAW0858	Peptide #7292 enco
26	8	2.8	55	22	AAW6451	Peptide #10488 enc
27	8	2.8	55	23	ABW45635	Human peptide enco
28	8	2.8	65	22	ABW43981	Peptide #11487 enc
29	8	2.8	65	22	ABW26890	Protein #8889 enco
30	8	2.8	65	22	AAW4984	Human brain expres
31	8	2.8	65	22	AAW7706	Human bone marrow
32	8	2.8	65	22	AAW21621	Peptide #8055 enco
33	8	2.8	65	22	AAW37923	Peptide #11960 enc
34	8	2.8	65	22	ABW46737	Human peptide enco
35	8	2.8	69	22	ABW03723	Novel human diagno
36	8	2.8	71	22	ABW03663	Novel human diagno
37	8	2.8	84	22	AAW61282	Propionibacterium
38	8	2.8	110	22	ABW20259	Novel human diagno
39	8	2.8	117	22	AAW08094	Human polypeptide
40	8	2.8	184	22	ABW67712	Drosophila melanog
41	8	2.8	190	23	ABW28041	Streptococcus poly
42	8	2.8	198	22	AAW23684	Human EST encoded
43	8	2.8	222	23	ABW41674	Human ovarian anti
44	8	2.8	286	22	ABW16279	Novel human diagno
45	8	2.8	307	22	ABW01186	Novel human diagno

## ALIGNMENTS

RESULT 1	
AAW47525	AAW47525 standard; Protein: 284 AA.
XX	AAW47525;
XX	21-JUL-1998 (first entry)
DT	
XX	Homo sapiens, tumour necrosis factor related ligand (TRELL).
DE	
XX	TRELL: tumour necrosis factor related ligand; tnfr; treatment;
KW	cancer; autoimmune disease; immune system; stimulation; suppression;
KW	graft rejection.
XX	
OS	Homo sapiens.
XX	
PN	WO9805783-A1.
PD	12-FEB-1998.
XX	
PF	07-AUG-1997; 97WO-US13945.
XX	
PR	18-MAR-1997; 97US-0040820.
PR	07-AUG-1996; 96US-0023541.
XX	18-OCT-1996; 96US-0028515.
PA	(BIOJ) BIOGEN INC.
PA	(URGE-) UNIV GENEVA FACULTY MEDICINE.
XX	
PI	Browning JL, Chicheportliche Y;
XX	
DR	WPI, 1998-145619/13.
DR	N-PSDB; AAW18600.



Query Match	Best Local Similarity	87.7%; Score 249;	DB 21;	Length 249;
Matches 249;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 36	MAARRSORRRGRGPERGTRALLVPLALGGLALACGLLLAVVSLGSRASLSAQERPAQETL 95			
DB 1	MAARRSORRRGRGPERGTRALLVPLALGGLALACGLLLAVVSLGSRASLSAQERPAQETL 60			
QY 96	VAEEPDODSESLNPQEEESODPAPFNLRLVPRPSAPKGRKTRARAIAAHYEVHPRPQGD 155			
DB 61	VAEEPDODSESLNPQEEESODPAPFNLRLVPRPSAPKGRKTRARAIAAHYEVHPRPQGD 120			
QY 156	GAQAGVDGTVSGWEABARINSSPLRYNRQIGEFIVTRAGLYLYLQCVHDEGKAAVYLKLD 215			
DB 121	GAQAGVDGTVSGWEABARINSSPLRYNRQIGEFIVTRAGLYLYLQCVHDEGKAAVYLKLD 180			
QY 216	LIVDVLALRCLEERSAPRAASSLGQGLRLCOVSSGLLALRPGSSLRRTLPMAHLKRAAPL 275			
DB 181	LIVDVLALRCLEERSAPRAASSLGQGLRLCOVSSGLLALRPGSSLRRTLPMAHLKRAAPL 240			
QY 276	TYFGLEFYVH 284			
DB 241	TYFGLEFYVH 249			
RESULT 4				
AA95338				
ID	AA95338 standard; Protein; 249 AA.			
AC	AA95338;			
XX				
DT	25-SEP-2000 (first entry)			
XX				
DE	Human PRO207 antitumour protein.			
XX				
KW	PRO207; human; antitumour; tumour; therapy; cytosolic; breast cancer; ovarian cancer; renal cancer; colorectal cancer; uterine cancer; prostate cancer; lung cancer; bladder cancer; central nervous system cancer; melanoma; leukaemia; neoplasm.			
XX				
OS	Homo sapiens.			
XX				
FT	Key	Location/Qualifiers		
FT	Peptide	1..40		
FT	Protein	/label= Signal_peptide		
FT		41..249		
FT		/label= PRO207		

FT	Modified-site	27..33	/note= "N-myristoylation"
FT	Modified-site	29..35	/note= "N-myristoylation"
FT	Modified-site	36..42	/note= "N-myristoylation"
FT	Modified-site	45..51	/note= "N-myristoylation"
FT	Modified-site	118..124	/note= "N-myristoylation"
FT	Modified-site	121..127	/note= "N-myristoylation"
FT	Modified-site	125..131	/note= "N-myristoylation"
FT	Modified-site	128..134	/note= "N-myristoylation"
FT	Modified-site	139..143	/note= "N-myristoylation"
FT	Modified-site	10..14	/note= "asn is N-glycosylated"
FT	Modified-site	97..101	/note= "amidation"
FT	Modified-site	24..35	/note= "amidation"
FT	Peptide		/note= "prokaryotic membrane lipoprotein lipid"
XX			
PN	WO200037638-A2.		
XX			
PD	29-JUN-2000.		
XX			
PP	02-DEC-1999;	99WO-US28565.	
XX			
PR	22-DEC-1998;	98US-0113296.	
PR	08-MAR-1999;	99WO-US05028.	
PR	21-APR-1999;	99US-0130232.	
PR	28-APR-1999;	99US-0131445.	
PR	14-MAY-1999;	99US-0134287.	
PR	20-JUL-1999;	98US-0144758.	
PR	26-JUL-1999;	99US-0145698.	
PR	15-SEP-1999;	99WO-US21090.	
PR	15-SEP-1999;	99WO-US21547.	
XX			
PA	(GETH ) GENENTECH INC.		
PI	Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;		
PI	Napier MA, Pitti RM, Wood WI;		
XX			
DR	WPI: 2000-442668/38.		
XX	N-PSDB: AAA49717.		
XX			
PT	Novel composition to inhibit neoplastic cell growth or for treating		
PT	tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,		
PT	PRO221, PRO224, PRO328, PRO301, PRO526, PRO356, PRO509 or		
PT	PRO866		
XX			
PS	Claim 19; Fig 4; 172pp: English.		
XX			
CC	The present sequence is that of human antitumour protein PRO207,		
CC	as deduced from a foetal kidney cDNA clone (see AAA49717). PRO207		
CC	shows amino acid sequence identity to tumour necrosis factor family		
CC	members, especially human lymphotoxin-beta (23.4%) and human CD40		
CC	ligand (19.8%). Mol.wt. is 27,216. A claimed method for inhibiting		
CC	the growth of a tumour cell comprises exposing the tumor cell		
CC	to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224, PRO301,		
CC	PRO526, PRO362, PRO356, PRO509 or PRO866 (see AA95337-49), their		
CC	agonists or chimeric polypeptides incorporating them. The tumour		
CC	is especially a cancer selected from breast, ovarian, renal,		
CC	colorectal, uterine, prostate, lung, bladder and central nervous		
CC	system cancer, melanoma and leukaemia. Methods for the recombinant		
CC	expression of the antitumour proteins are also provided.		
XX			
Sequence	249 AA;		

Query Match 87.7%; Score 249; DB 21; Length 249;





DR WPI: 1998-447255/38.  
 DR N-PSDB: AAV47613.  
 XX  
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
 PT treatment of autoimmune disease, tumours and inflammation  
 XX  
 PS Claim 16; Page 123-4; 142pp; English.  
 XX  
 CC The TNF-related endothelium proliferative agent (TREPA), or its  
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
 CC promote wound healing or tissue grafting, by promoting vascularisation,  
 CC also to induce apoptosis for treating cancer and eliminating autoreactive  
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
 CC TREPA peptides can also be used to target cytotoxic agents or for  
 CC affinity isolation of the corresponding receptor, the nucleic acid for  
 CC which can be used to transform tumour cells to render them more  
 CC responsive to TREPA and to screen for TREPA mimics.  
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
 CC vascularisation), inflammation or a wide range of autoimmune conditions,  
 CC conditions involving abnormal stimulation of epithelial cells (e.g.  
 CC atherosclerosis), for birth control (inhibiting ovulation and placental  
 CC formation) or other angiogenic conditions (e.g. ulcers).  
 CC  
 SQ Sequence 249 AA:  
 Query Match 84.9%; Score 241; DB 19; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-213;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 44 RRGRRGEPGTTALVPLALGIGLALACGLLAVVSGSRASISAOPAEELVAEEDDP 103  
 DB 9 RRGRRGEPGTTALVPLALGIGLALACGLLAVVSGSRASISAOPAEELVAEEDDP 68  
 QY 104 SELNPQTEESODPAPFLNVLVRRPSAPKGRKTRARRAIAHYEVHPRGOGAAGVGD 163  
 DB 69 SELNPQTEESODPAPFLNVLVRRPSAPKGRKTRARRAIAHYEVHPRGOGAAGVGD 128  
 QY 164 TVSGWEARINSSPLRYRQIGEFIVTRAGIYLYCQVHFDEGKAVYIKLDLVDGYLA 223  
 DB 129 TVSGWEARINSSPLRYRQIGEFIVTRAGIYLYCQVHFDEGKAVYIKLDLVDGYLA 188  
 QY 224 LRCLFEFSATTAASSIGPOLRLCOVSGLLALRPSSLRITLPMHAKAPFLTYFGGLV 283  
 DB 189 LRCLFEFSATTAASSIGPOLRLCOVSGLLALRPSSLRITLPMHAKAPFLTYFGGLV 248  
 QY 284 H 284  
 DB 249 H 249  
 RESULT 7  
 AAE00891  
 ID AAE00891 standard; Protein: 249 AA.  
 XX  
 AC AAE00891;  
 XX  
 DT 04-JUL-2001 (first entry)  
 XX  
 DE Human TREPA (TNF related endothelium proliferative agent).  
 XX  
 KW Human: tumour necrosis factor; TNF; angiogenesis; wound healing;  
 KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;  
 KW grafting; vulnerability.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 98..249  
 FT /label= Extracellular\_domain  
 XX  
 PN US6207642-B1.  
 XX

PD 27-MAR-2001.  
 XX  
 PF 26-JUN-1998; 98US-0105343.  
 XX  
 PR 12-FEB-1997; 97US-0798692.  
 PR 10-FEB-1998; 98US-0021706.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Wiley SR;  
 XX  
 DR WPI: 2001-280760/29.  
 DR N-PSDB: AAD04350.  
 XX  
 PT Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 XX  
 PS Claim 1; Column 75-76; 53pp; English.  
 XX  
 CC The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts.  
 CC The present amino acid sequence is clone ID #690050 human TREPA.  
 CC  
 SQ Sequence 249 AA:  
 Query Match 84.9%; Score 241; DB 22; Length 249;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-213;  
 Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 44 RRGRRGEPGTTALVPLALGIGLALACGLLAVVSGSRASISAOPAEELVAEEDDP 103  
 DB 9 RRGRRGEPGTTALVPLALGIGLALACGLLAVVSGSRASISAOPAEELVAEEDDP 68  
 QY 104 SELNPQTEESODPAPFLNVLVRRPSAPKGRKTRARRAIAHYEVHPRGOGAAGVGD 163  
 DB 69 SELNPQTEESODPAPFLNVLVRRPSAPKGRKTRARRAIAHYEVHPRGOGAAGVGD 128  
 QY 164 TVSGWEARINSSPLRYRQIGEFIVTRAGIYLYCQVHFDEGKAVYIKLDLVDGYLA 223  
 DB 129 TVSGWEARINSSPLRYRQIGEFIVTRAGIYLYCQVHFDEGKAVYIKLDLVDGYLA 188  
 QY 224 LRCLFEFSATTAASSIGPOLRLCOVSGLLALRPSSLRITLPMHAKAPFLTYFGGLV 283  
 DB 189 LRCLFEFSATTAASSIGPOLRLCOVSGLLALRPSSLRITLPMHAKAPFLTYFGGLV 248  
 QY 284 H 284  
 DB 249 H 249  
 RESULT 8  
 AAU03499  
 ID AAU03499 standard; Protein: 273 AA.  
 XX  
 AC AAU03499;  
 XX  
 DT 26-SEP-2001 (first entry)  
 XX  
 DE TWEAK extracellular domain-containing fusion protein.  
 XX  
 KW TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;  
 KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;  
 KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;  
 KW rubeosis; uveitis; macular degeneration; arthritis; rheumatism;  
 KW corneal graft neovascularisation; psoriasis; metastatic condition;  
 KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;  
 KW

KW preneoplastic condition; myocardial angiogenesis; wound granulation;  
 KW scleroderma; vascular adhesion; telangiectasia; ischemia; human;  
 KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;  
 KW peripheral atherosclerosis; pDC409-L2-TWEAK; TWEAK receptor; TWEAKR;  
 KW fusion protein.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN MO200145730-A2.  
 XX  
 PD 28-JUN-2001.  
 XX  
 PF 19-DEC-2000; 2000MO-US34755.  
 XX  
 PR 20-DEC-1999; 99US-0172878.  
 PR 10-MAY-2000; 2000US-0203347.  
 XX  
 PA (IMM) ) IMMUNEX CORP.  
 XX  
 PI Wiley SR;  
 XX  
 DR WPI: 2001-417975/44.  
 DR N-PSDB: AAS03964.  
 XX  
 XX  
 PT Modulating angiogenesis in a mammal for treating diseases mediated by  
 PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or  
 PT peripheral tissue, by administering antagonist or agonist of TWEAK  
 PT receptor  
 XX  
 PS Example 1; Page 41; 46pp; English.  
 XX  
 CC The sequence represents a fusion protein encoded by the  
 CC expression vector pDC409-L2-TWEAK. The fusion protein comprises a  
 CC growth hormone leader, a leucine zipper multimerisation domain, and  
 CC the extracellular domain of human TWEAK. The fusion protein was  
 CC used in the isolation of human TWEAK receptor (TWEAKR)-expressing  
 CC clones from a COS cell human cDNA library. The TWEAK protein is  
 CC a member of the tumour necrosis factor (TNF) family and induces  
 CC angiogenesis. TWEAKR may therefore be used to screen for and  
 CC develop TWEAKR agonists and antagonists for the modulation of  
 CC angiogenesis, to be used in the treatment and diagnosis of human disease.  
 CC The disorders mediated by angiogenesis include ocular disorders  
 CC characterised by ocular neovascularisation such as diabetic retinopathy,  
 CC neovascular glaucoma, retinoblastoma, retinopathy of prematurity,  
 CC retrolental fibroplasia, rubeosis, uveitis, macular degeneration and  
 CC corneal graft neovascularisation, and inflammatory diseases such as  
 CC arthritis, rheumatism and psoriasis. Other treatable diseases include  
 CC malignant and metastatic conditions such as sarcomas and carcinomas,  
 CC benign tumours and preneoplastic conditions, myocardial angiogenesis,  
 CC hemophilic joints, scleroderma, vascular adhesions, atherosclerotic  
 CC plaque neovascularisation, telangiectasia, wound granulation, coronary  
 CC atherosclerosis, peripheral atherosclerosis and ischaemia.  
 CC  
 XX  
 SQ Sequence 273 AA;  
 XX  
 Query Match 72.9%; Score 207; DB 22; Length 273;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-181;  
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 78 SLGRASLSLAQEPQAEELVAEEDDPSELNPTQEESSODPAPFLRLVLRPRRSAPKGRKTR 137  
 DB 67 SLGRASLSLAQEPQAEELVAEEDDPSELNPTQEESSODPAPFLRLVLRPRRSAPKGRKTR 126  
 QY 138 ARRAIAAHYEHPRPGDGAAGVDTGVSQWEARINSSPLRNRQIGEIFYTRAGILYTL 197  
 DB 127 ARRAIAAHYEHPRPGDGAAGVDTGVSQWEARINSSPLRNRQIGEIFYTRAGILYTL 186  
 QY 198 LYCOVHDEGKAVYLIKIDLLVDGVALRCLFEFSATASSLGPOLRLCOVSGLLALRPGS 257  
 DB 187 LYCOVHDEGKAVYLIKIDLLVDGVALRCLFEFSATASSLGPOLRLCOVSGLLALRPGS 246  
 QY 258 SLRITLPMAHILKAAPFLTYGFLFOVH 284

DB 247 SLRITLPMAHILKAAPFLTYGFLFOVH 273  
 RESULT 9  
 AAE00895  
 ID AAE00895 standard; Protein; 146 AA.  
 XX  
 AC AAE00895;  
 XX  
 DT 04-JUL-2001 (first entry)  
 XX  
 DE Human TREPA (TNF related endothelium proliferative agent) fragment.  
 XX  
 KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing;  
 KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;  
 KW grafting; vulneryary.  
 XX  
 OS Homo sapiens.  
 OS  
 PN US6207642-B1.  
 PN  
 PD 27-MAR-2001.  
 XX  
 PF 26-JUN-1998; 98US-0105343.  
 XX  
 PR 12-FEB-1997; 97US-0798692.  
 PR 10-FEB-1998; 98US-0021706.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Wiley SR;  
 XX  
 DR WPI: 2001-280760/29.  
 XX  
 PT Inducing angiogenesis in mammal at desired sites for promoting wound  
 PT healing, by administering soluble fragment of extracellular domain of  
 PT tumor necrosis factor related endothelium proliferative agent protein  
 PT  
 XX  
 PS Example 14; Fig 1; 53pp; English.  
 XX  
 CC The present invention relates to extracellular signal molecules,  
 CC particularly members of tumour necrosis factor (TNF) family molecules  
 CC designated as TREPA (TNF related endothelium proliferative agent).  
 CC Soluble biologically active TREPA are used to treat TREPA-associated  
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
 CC in human for promoting wound healing and for vascularising grafted  
 CC tissue for successful grafting and to promote tissue grafts.  
 CC The present amino acid sequence is human TREPA fragment.  
 CC  
 XX  
 SQ Sequence 146 AA;  
 XX  
 Query Match 51.4%; Score 146; DB 22; Length 146;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-126;  
 Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 139 RRAIAAHYEHPRPGDGAAGVDTGVSQWEARINSSPLRNRQIGEIFYTRAGILYTL 198  
 DB 1 RRAIAAHYEHPRPGDGAAGVDTGVSQWEARINSSPLRNRQIGEIFYTRAGILYTL 60  
 QY 199 YCOVHDEGKAVYLIKIDLLVDGVALRCLFEFSATASSLGPOLRLCOVSGLLALRPGS 258  
 DB 61 YCOVHDEGKAVYLIKIDLLVDGVALRCLFEFSATASSLGPOLRLCOVSGLLALRPGS 120  
 QY 259 LRITLPMAHILKAAPFLTYGFLFOVH 284  
 DB 121 LRITLPMAHILKAAPFLTYGFLFOVH 146  
 RESULT 10  
 AAM29746  
 ID AAM29746 standard; Protein; 189 AA.

XX AAM29746;  
AC  
XX 27-OCT-1998 (first entry)  
DT  
XX  
XX  
DE TNF related endothelium proliferative agent protein 2.  
DE  
XX  
XX TNF; endothelium proliferative agent; TREPA; wound healing; cancer;  
KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.  
XX  
OS Homo sapiens.  
XX  
PN W09835061-A2.  
XX  
PD 13-AUG-1998.  
XX  
PF 12-FEB-1998; 98WO-US02859.  
XX  
PR 10-FEB-1998; 98US-0021706.  
PR 12-FEB-1997; 97US-0798692.  
XX  
XX (ABBO ) ABBOTT LAB.  
PA  
XX  
PI Wiley SR;  
XX  
DR WPI; 1998-447255/38.  
XX  
PT Detecting nucleic acid encoding TREPA - useful for diagnosis and  
PT treatment of autoimmune disease, tumours and inflammation  
XX  
PS Claim 16; Page 125-6; 142pp; English.  
XX  
CC The TNF-related endothelium proliferative agent (TREPA), or its  
CC activators or agonists, are used to treat a deficit of TREPA, e.g. to  
CC promote wound healing or tissue grafting, by promoting vascularisation,  
CC also to induce apoptosis for treating cancer and eliminating autoreactive  
CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.  
CC TREPA peptides can also be used to target cytotoxic agents or for  
CC affinity isolation of the corresponding receptor, the nucleic acid for  
CC which can be used to transform tumour cells to render them more  
CC responsive to TREPA and to screen for TREPA mimics.  
CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat  
CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting  
CC vascularisation), inflammation or a wide range of autoimmune conditions,  
CC conditions involving abnormal stimulation of epithelial cells (e.g.  
CC atherosclerosis), for birth control (inhibiting ovulation and placental  
CC formation) or other angiogenic conditions (e.g. ulcers).  
XX  
SQ Sequence 189 AA;  
Query Match 50.4%; Score 143; DB 19; Length 189;  
Best Local Similarity 100.0%; Pred. No. 4.6e-123;  
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 142 IAAHEVHPRGQDAQAGVDGTVSGWEARINSSPLRYNRQIGFIVTRAGLYLYCQ 201  
DB 47 IAAHEVHPRGQDAQAGVDGTVSGWEARINSSPLRYNRQIGFIVTRAGLYLYCQ 106  
QY 202 VHFDEGKAVYTKLDLVGVYALRCLEESATPAASLGPOLRLCOVSGLLALRPSSIRI 261  
DB 107 VHFDEGKAVYTKLDLVGVYALRCLEESATPAASLGPOLRLCOVSGLLALRPSSIRI 166  
QY 262 RTLPMAHKAAAPFLTYFGLFOVH 284  
DB 167 RTLPMAHKAAAPFLTYFGLFOVH 189

RESULT 11  
AAE00892  
ID AAE00892 standard; Protein: 189 AA.  
XX  
XX AAE00892;  
XX

DT 04-JUL-2001 (first entry)  
XX  
XX Human ULAflag TREPA soluble construct.  
DE  
XX  
XX Human; tumour necrosis factor; TNF; angiogenesis; wound healing; tumour;  
KW TREPA; TNF related endothelium proliferative agent; metastasis; grafting;  
KW vulnerable; HUVEC; human umbilical vein endothelial cell; ULAflag.  
XX  
OS Homo sapiens.  
XX  
PN US6207642-B1.  
XX  
PD 27-MAR-2001.  
XX  
PF 26-JUN-1998; 98US-0105343.  
XX  
PR 12-FEB-1997; 97US-0798692.  
PR 10-FEB-1998; 98US-0021706.  
XX  
XX (ABBO ) ABBOTT LAB.  
PA  
XX  
PI Wiley SR;  
XX  
DR WPI; 2001-280760/29.  
XX  
PT Inducing angiogenesis in mammal at desired sites for promoting wound  
PT healing, by administering soluble fragment of extracellular domain of  
PT tumor necrosis factor related endothelium proliferative agent protein  
XX  
XX Example 2; Column 75-78; 53pp; English.  
XX  
PS The present invention relates to extracellular signal molecules,  
XX particularly members of tumour necrosis factor (TNF) family molecules  
CC designated as TREPA (TNF related endothelium proliferative agent).  
CC Soluble biologically active TREPA are used to treat TREPA-associated  
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis  
CC in human for promoting wound healing and for vascularising grafted  
CC tissue for successful grafting and to promote tissue grafts. The present  
CC amino acid sequence is human ULAflag TREPA soluble construct. This  
CC sequence which is a biologically active molecule is capable of inducing  
CC proliferation in HUVEC (human umbilical vein endothelial cells).  
XX  
SQ Sequence 189 AA;  
Query Match 50.4%; Score 143; DB 22; Length 189;  
Best Local Similarity 100.0%; Pred. No. 4.6e-123;  
Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 142 IAAHEVHPRGQDAQAGVDGTVSGWEARINSSPLRYNRQIGFIVTRAGLYLYCQ 201  
DB 47 IAAHEVHPRGQDAQAGVDGTVSGWEARINSSPLRYNRQIGFIVTRAGLYLYCQ 106  
QY 202 VHFDEGKAVYTKLDLVGVYALRCLEESATPAASLGPOLRLCOVSGLLALRPSSIRI 261  
DB 107 VHFDEGKAVYTKLDLVGVYALRCLEESATPAASLGPOLRLCOVSGLLALRPSSIRI 166  
QY 262 RTLPMAHKAAAPFLTYFGLFOVH 284  
DB 167 RTLPMAHKAAAPFLTYFGLFOVH 189

RESULT 12  
AAM93590  
ID AAM93590 standard; Protein: 208 AA.  
XX  
XX AAM93590;  
XX  
XX 18-JUN-1999 (first entry)  
DT  
XX  
XX Human TNRL3 protein.  
DE  
XX  
KW Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;

KW developmental abnormality; gestational abnormality; prostate cancer;  
 KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW apoptosis; human.  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 PN W09911791-A2.  
 PD 11-MAR-1999.  
 XX  
 PF 04-SEP-1998; 98WC-US18393.  
 XX  
 PR 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 PI Chaudhary PM;  
 XX  
 XX WPI; 1999-205191/17.  
 DR N-PSDB; AAX23424.  
 XX  
 XX New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 XX  
 PS Claim 40; Fig 13A; 156pp; English.  
 XX  
 XX This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC moiety. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/ active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the changer in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active  
 CC fragments and APO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
 CC activity. The method is performed in vivo or in vitro. APO polypeptides  
 CC are all useful as immunogens for preparing antibodies. APO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 XX  
 XX Sequence 208 AA;  
 SQ  
 Query Match 37.7%; Score 107; DB 20; Length 208;  
 Best Local Similarity 99.5%; Pred. No. 5.8e-90;  
 Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 77 VSLSRSASLSAQEPQAEELVAEEDQDPSELNPTQESODAPAPLNLVRRSPKGRKT 136  
 DB 1 VSLSRSASLSAQEPQAEELVAEEDQDPSELNPTQESODAPAPLNLVRRSPKGRKT 60  
 QY 137 RARAIAAHYEVHPRPODQAGVDSGSGWEARINSSPLLYNQIGFTYRAGLY 196  
 DB 61 RARAIAAHYEVHPRPODQAGVDSGSGWEARINSSPLLYNQIGFTYRAGLY 120  
 QY 197 YLYCOVHFDEGKAYYLRDLVDVLAIRCLIEFSATASLGPQLRCQVSGLLALRP 256  
 DB 121 YLYCOVHFDEGKAYYLRDLVDVLAIRCLIEFSATASLGPQLRCQVSGLLALRP 180  
 QY 257 SSLRIRLPMWHLKAAPFLTYFGLFOVH 284  
 DB 181 SSLRIRLPMWHLKAAPFLTYFGLFOVH 208

RESULT 13  
 AAW93591  
 ID AAW93591 standard; Protein: 211 AA.  
 XX  
 AC AAW93591;  
 XX  
 DT 18-JUN-1999 (first entry)  
 XX  
 DE Mouse TNRL3 protein.  
 XX  
 KW Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;  
 KW developmental abnormality; gestational abnormality; prostate cancer;  
 KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;  
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;  
 KW apoptosis; mouse.  
 XX  
 OS Mus sp.  
 XX  
 XX W09911791-A2.  
 PN 11-MAR-1999.  
 PD 11-MAR-1999.  
 XX  
 PF 04-SEP-1998; 98WC-US18393.  
 XX  
 PR 05-SEP-1997; 97US-0924634.  
 XX  
 PA (UNIW ) UNIV WASHINGTON.  
 PI Chaudhary PM;  
 XX  
 XX WPI; 1999-205191/17.  
 DR N-PSDB; AAX23425.  
 XX  
 XX New Tumor Necrosis Factor family receptor polypeptides and ligands -  
 PT useful for diagnosis and treatment of prostate cancer and  
 PT developmental or gestational abnormalities  
 XX  
 PS Claim 40; Fig 13B; 156pp; English.  
 XX  
 XX This invention describes isolated Tumor Necrosis Factor (TNF) family  
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active  
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or  
 CC their active fragments. APO4 is useful for diagnosing prostate cancer  
 CC by determining levels of APO4 in an individual. Prostate cancer can also  
 CC be treated using APO4 selective binding agents linked to a therapeutic  
 CC moiety. APO4 polypeptides are also useful for identifying selective  
 CC binding agents, useful in diagnosis/treatment of disease by binding of  
 CC agents to the polypeptide/active fragment which is extracellular, or  
 CC expressed on the cell surface. The binding is preferably performed in  
 CC vivo. APO4 polypeptides/ active fragments are also useful for screening  
 CC for agonists and antagonists by binding and observing the changer in APO4  
 CC activity. Effective pharmacological agents useful in diagnosis or  
 CC treatment of disease are also identified using APO4 polypeptides/active  
 CC fragments and APO4 signal transducer molecules that specifically interact  
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4  
 CC activity. The method is performed in vivo or in vitro. APO polypeptides  
 CC are all useful as immunogens for preparing antibodies. APO4 is also  
 CC useful for diagnosis/treatment of developmental or gestational  
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line  
 CC MCF-7, and induced apoptosis.  
 XX  
 XX Sequence 211 AA;  
 SQ  
 Query Match 16.2%; Score 46; DB 20; Length 211;  
 Best Local Similarity 100.0%; Pred. No. 6.1e-34;  
 Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 239 GPQLRCQVSGLLALRPSSSLRIRLPMWHLKAAPFLTYFGLFOVH 284  
 DB 166 GPQLRCQVSGLLALRPSSSLRIRLPMWHLKAAPFLTYFGLFOVH 211

RESULT 14

AAW47524  
ID AAW47524 standard; Protein: 225 AA.  
XX  
AC AAW47524;  
XX  
DT 21-JUL-1998 (first entry)  
XX  
DE Mus musculus tumour necrosis factor related ligand (TRELL).  
XX  
KW TRELL; tumour necrosis factor related ligand; tnfr; treatment;  
KM cancer; autoimmune disease; immune system; stimulation; suppression;  
XX graft rejection.  
XX  
OS Mus musculus.  
XX  
FH Key Location/Qualifiers  
FT Domain 1..21  
FT /note= "hydrophobic, transmembrane domain"  
XX  
PN WO9805783-A1.  
XX  
PD 12-FEB-1998.  
XX  
PF 07-AUG-1997; 97WO-US13945.  
XX  
PR 18-MAR-1997; 97US-0040820.  
PR 07-AUG-1996; 96US-0023541.  
PR 18-OCT-1996; 96US-0028515.  
XX  
PA (BIOJ ) BIOGEN INC.  
PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.  
XX  
PI Browning JL, Chicheportriche Y;  
XX  
DR WPI: 1998-145619/13.  
DR N-PSDB; AAV18599.  
XX  
PT Tumour necrosis factor related ligand - useful for, e.g. treating  
PT cancer, auto-immune disease and immune responses to tissue grafts  
XX  
PS Claim 12; Pages 48-50; 69pp; English.  
XX  
XX The sequence is that of mouse tumour necrosis factor related  
CC ligand (TRELL). TRELL or active fragments can be included with a  
CC carrier in pharmaceutical compositions to treat cancer, autoimmune  
CC diseases or immune responses to tissue grafts, or to stimulate or  
CC suppress the immune system. It is useful to screen for TRELL  
CC receptors, by labelling with a detectable label and screening  
CC compositions for binding. Agents interfering with TRELL-receptor  
CC binding can also be screened for, can then be administered,  
CC optionally with interferon- gamma, to induce cell death or  
CC treat, suppress or alter immune responses (especially involving human  
CC adenocarcinoma cells) involving a signal pathway between TRELL and its  
CC receptor. It's coding sequence can be used in gene therapy for  
CC TRELL-related disorders in mammals (especially humans), e.g. tumours,  
CC autoimmune and inflammatory diseases or inherited genetic disorders,  
CC by introducing into cells, and expressing, therapeutically effective  
CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.  
CC It may also be of use in the preparation of prepare probes for  
CC screening natural/synthetic DNAs for TRELL-encoding sequences  
CC and for antisense therapy.  
XX  
SQ Sequence 225 AA:  
  
Query Match 11.3%; Score 32; DB 19; Length 225;  
Best Local Similarity 100.0%; Pred. No. 4.7e-21;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 15  
AAB07527  
ID AAB07527 standard; protein: 225 AA.  
XX  
AC AAB07527;  
XX  
DT 20-OCT-2000 (first entry)  
XX  
DE Amino acid sequence of a soluble recombinant murine TWEAK protein.  
XX  
KW TWEAK protein; immunological disorder; immune response; inflammation;  
KM TWEAK blocking agent; autoimmune disease; organ transplant rejection;  
XX Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock; tumour.  
XX  
OS Mus sp.  
XX  
PN WO200042073-A1.  
XX  
PD 20-JUL-2000.  
XX  
PF 14-JAN-2000; 2000WO-US01044.  
XX  
PR 15-JAN-1999; 99US-0116168.  
XX  
PA (BIOJ ) BIOGEN INC.  
XX  
PI Remmert P;  
XX  
DR WPI: 2000-476036/41.  
XX  
PT Preventing and treating immune responses using modulators, especially  
PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for  
PT treating e.g. inflammation and graft versus host disease -  
XX  
PS Disclosure: Fig 1; 45pp; English.  
XX  
XX The present sequence represents a TWEAK protein. The specification  
CC describes a method for preventing or treating an immunological  
CC disorder and/or inhibiting an immune response in an animal. The  
CC method comprises administering a TWEAK blocking agent. The method may  
CC be used for preventing and treating immune disorders associated with  
CC inappropriate expression and/or activity of TWEAK. These disorders  
CC include autoimmune diseases, acute and chronic inflammation, organ  
CC transplant rejection, Graft-versus-Host disease (GVHD), lymphoid cell  
CC malignancies, septic and other forms of shock, loss of immune  
CC responsiveness (as seen in human immunodeficiency virus (HIV)  
CC infections) and failure of the immune response to tumour growth.  
XX  
SQ Sequence 225 AA:  
  
Query Match 11.3%; Score 32; DB 21; Length 225;  
Best Local Similarity 100.0%; Pred. No. 4.7e-21;  
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

139 RRAIAHYEVHPRPGDGAQAGVDGTVSGMEE 170  
|||||  
80 RRAIAHYEVHPRPGDGAQAGVDGTVSGMEE 111

RESULT 16  
AAG01265  
ID AAG01265 standard; Protein: 58 AA.  
XX  
AC AAG01265;  
XX  
DT 06-OCT-2000 (first entry)  
XX  
DE Human secreted protein, SEQ ID NO: 5346.  
XX  
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;  
KM gene therapy; chromosome mapping.  
XX  
OS Homo sapiens.

```

XX  EP1033401-A2.
PN  06-SEP-2000.
PD  21-FEB-2000; 2000EP-0200610.
XX  26-FEB-1999; 99US-0122487.
XX  (GENSET) GENSET.
PA  Dumas Milne Edwards J, Duclert A, Giordano J;
XX  WPI: 2000-500381/45.
DR  N-PSDB: AAC01271.
XX  New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT  obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT  diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX  Claim 13; SEQ ID 5346; 71pp + CD-ROM; English.
PS
XX  The present sequence is a polypeptide encoded by one of a large number
CC  of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC  were prepared from total human RNAs or polyA+ RNAs derived from 30
CC  different tissues. EST sequences usually correspond mainly to the 3'
CC  untranslated region (UTR) of the mRNA because they are often obtained
CC  from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC  isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC  those cases where longer cDNA sequences have been obtained, the full 5'
CC  UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC  ends and can therefore be used to obtain full length cDNAs and genomic
CC  DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC  chromosome mapping procedures. They are used to obtain upstream
CC  regulatory sequences and to design expression and secretion vectors.
SQ  Sequence 58 AA:
Query Match 3.2%; Score 9; DB 21; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSLDFEIS 9
Db 46 MSLDFEIS 54
RESULT 17
AAAG01266
ID AAC01266 standard; Protein: 58 AA.
AC
XX  AAG01266;
XX  06-OCT-2000 (first entry)
DT
XX  Human secreted protein, SEQ ID NO: 5347.
DE
XX  Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW  gene therapy; chromosome mapping.
XX
XX  Homo sapiens.
OS
XX  EPI033401-A2.
PN
XX  06-SEP-2000.
PD
XX  21-FEB-2000; 2000EP-0200610.
PF
XX  26-FEB-1999; 99US-0122487.
PR
XX  (GENSET) GENSET.
PA
XX  Dumas Milne Edwards J, Duclert A, Giordano J;
PI

```

```

XX  WPI: 2000-500381/45.
DR  N-PSDB: AAC01272.
XX  New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT  obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
PT  diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX  Claim 13; SEQ ID 5347; 71pp + CD-ROM; English.
PS
XX  The present sequence is a polypeptide encoded by one of a large number
CC  of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC  were prepared from total human RNAs or polyA+ RNAs derived from 30
CC  different tissues. EST sequences usually correspond mainly to the 3'
CC  untranslated region (UTR) of the mRNA because they are often obtained
CC  from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC  isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC  those cases where longer cDNA sequences have been obtained, the full 5'
CC  UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC  ends and can therefore be used to obtain full length cDNAs and genomic
CC  DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC  chromosome mapping procedures. They are used to obtain upstream
CC  regulatory sequences and to design expression and secretion vectors.
SQ  Sequence 58 AA:
Query Match 3.2%; Score 9; DB 21; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSLDFEIS 9
Db 46 MSLDFEIS 54
RESULT 18
AAV14906
ID AAV14906 standard; Protein: 748 AA.
AC
XX  AAV14906;
XX  25-OCT-1999 (first entry)
DT
XX  Extended amino acid sequence for GVA-9.
DE
XX  Mycobacterium vaccae protein; antigen; T cell activation; cytokine;
KW  dendritic cell maturation; infectious disease; immune disorder; cancer;
KW  respiratory system; mycobacterial infection; allergy; tuberculosis;
KW  leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;
KW  dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
KW  squamous cell carcinoma; melanoma.
XX
XX  Mycobacterium vaccae.
OS
XX  WO9932634-A2.
PN
XX  01-JUL-1999.
PD
XX  23-DEC-1998; 98WO-NZ00189.
PF
XX  04-DEC-1998; 98US-0205426.
PR  23-DEC-1997; 97US-0996624.
PR  23-DEC-1997; 97US-0997080.
PR  23-DEC-1997; 97US-0997362.
PR  11-JUN-1998; 98US-0095855.
PR  17-SEP-1998; 98US-0156181.
XX
XX  (GENE-) GENESIS RES & DEV CORP LTD.
PA  Prestidge RL, Skinner MA, Tan P, Visser ES, Watson J;
PI  WPI: 1999-430163/36.
DR  N-PSDB: AAZ11368.

```

XX Enhancing immune response to an antigen  
 PT  
 PS Claim 1: Page 209-210; 243pp; English.  
 XX  
 CC The invention provides heat-killed *Mycobacterium vaccae*, or recombinant  
 CC *M. vaccae* proteins. The *M. vaccae* proteins may be employed to activate  
 CC T cells and natural killer cells, to stimulate the production of  
 CC cytokines, to enhance the expression of co-stimulatory molecules on  
 CC dendritic cells and monocytes, and to enhance dendritic cell maturation  
 CC and function. The proteins can be expressed by standard recombinant  
 CC methodology. Pharmaceutical compositions comprising the proteins or  
 CC nucleic acid sequences encoding the proteins can be used for the  
 CC treatment, prevention, and detection of disorders including infectious  
 CC diseases, immune disorders and cancer. In particular, the compounds and  
 CC methods are used for treatment of diseases of the respiratory system,  
 CC such as mycobacterial infections, asthma, allergies, tuberculosis,  
 CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as  
 CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,  
 CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell  
 CC carcinoma and melanoma.  
 SQ Sequence 748 AA;  
 Query Match 3.2%; Score 9; DB 20; Length 748;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 60 ALGLGIALA 68  
 DB 282 ALGLGIALA 290  
 |||||  
 RESULT 19  
 ABB73512  
 ID ABB73512 standard; Protein; 749 AA.  
 AC ABB73512;  
 XX  
 XX 08-APR-2002 (first entry)  
 DE M vaccae GVS-9 protein SEQ ID NO: 154.  
 XX  
 XX Skin disorder: psoriasis; atopic dermatitis; allergic contact dermatitis;  
 KW alopecia areata; skin cancer; *Mycobacterium vaccae*; melanoma; cytostatic;  
 KW antipsoriatic; dermatological; antiinflammatory; antiallergic;  
 KW Th2 immune response; immunomodulatory.  
 XX  
 OS *Mycobacterium vaccae*.  
 XX  
 PN US6328978-B1.  
 XX  
 PD 11-DEC-2001.  
 XX  
 PF 02-JUN-1999; 99US-0324542.  
 XX  
 PR 23-DEC-1997; 97US-0997080.  
 XX  
 PA (GENE-) GENESIS RES & DEV CORP LTD.  
 XX  
 PI Watson JD, Tan PLJ, Prestidge R;  
 XX  
 DR WPI: 2002-138361/18.  
 DR N-PSDB: ABL36274.  
 XX  
 XX Inhibiting skin inflammation associated with skin disorder e.g.  
 PT psoriasis, by administering composition comprising delipidated and  
 PT deglycolipidated *Mycobacterium vaccae* cells or *Mycobacterium vaccae*  
 PT culture filtrate  
 XX  
 PS Example 6; Column 161-164; 116pp; English.  
 CC The present invention relates to a method of inhibiting skin inflammation

CC associated with a skin disorder selected from psoriasis, atopic  
 CC dermatitis and allergic contact dermatitis, which involves administering  
 CC a composition containing delipidated and deglycolipidated *Mycobacterium*  
 CC *vaccae* cells or *M. vaccae* culture filtrate. The skin disorder to be  
 CC treated may also include alopecia areata, and skin cancers such as basal  
 CC cell carcinoma, squamous cell carcinoma and melanoma. The composition  
 CC acts by inhibiting the Th2 immune response. The present sequence is a  
 CC protein described in the exemplification of the invention.  
 SQ Sequence 749 AA;  
 Query Match 3.2%; Score 9; DB 23; Length 749;  
 Best Local Similarity 100.0%; Pred. No. 20;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 60 ALGLGIALA 68  
 DB 282 ALGLGIALA 290  
 |||||  
 RESULT 20  
 AAU51863  
 ID AAU51863 standard; Protein; 54 AA.  
 AC AAU51863;  
 XX  
 XX 27-FEB-2002 (first entry)  
 DE Propionibacterium acnes immunogenic protein #12759.  
 XX  
 XX SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;  
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
 KW dermatological; osteopathic; neuroprotectant.  
 XX  
 OS Propionibacterium acnes.  
 XX  
 PN WO200181581-A2.  
 XX  
 PD 01-NOV-2001.  
 XX  
 PF 20-APR-2001; 2001WO-US12865.  
 XX  
 PR 21-APR-2000; 2000US-199047P.  
 PR 02-JUN-2000; 2000US-208841P.  
 PR 07-JUL-2000; 2000US-216747P.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
 XX  
 DR WPI: 2001-616774/71.  
 DR N-PSDB: AAS59552.  
 XX  
 XX Propionibacterium acnes polypeptides and nucleic acids useful for  
 PT vaccinating against and diagnosing infections, especially useful for  
 PT treating acne vulgaris -  
 XX  
 PS Example 1; SEQ ID NO 13058; 1069pp; English.  
 XX  
 CC Sequences AAU39105-AAU66017 represent Propionibacterium acnes immunogenic  
 CC polypeptides. The proteins and their associated DNA sequences are used in  
 CC the treatment, prevention and diagnosis of medical conditions caused by  
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.  
 CC P. acnes is also involved in infections of bone, joints and the central  
 CC nervous system, however it is particularly involved in the inflammatory  
 CC lesions associated with acne vulgaris. A method for detecting the  
 CC presence or absence of P. acnes in a patient comprises contacting a  
 CC sample with a binding agent that binds to the proteins of the invention  
 CC and determining the amount of bound protein in the sample. The  
 CC polypeptides may be used as antigens in the production of antibodies

CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 54 AA:

Query Match 2.88; Score 8; DB 22; Length 54;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 RRLPLPRS 18  
|  
Db 21 RRLPLPRS 28

## RESULT 21

ABBA2637  
ID ABB42637 standard; Peptide: 55 AA.

XX ABB42637;

DT 04-FEB-2002 (first entry)

DE Peptide #10143 encoded by human foetal liver single exon probe.

KW Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

PN W0200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human foetal liver.

PS Claim 27; SEQ ID NO 35272; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC foetal liver. The present sequence is a peptide encoded by a single exon

CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 55 AA:

Query Match 2.88; Score 8; DB 22; Length 55;  
Best Local Similarity 100.0%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGIALA 68  
|  
Db 12 LGIGIALA 19

## RESULT 22

ABB25983  
ID ABB25983 standard; Protein: 55 AA.

XX ABB25983;

DT 23-JAN-2002 (first entry)

DE Protein #7982 encoded by probe for measuring heart cell gene expression.

KW Human; gene expression; heart; microarray; vascular system;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

OS Homo sapiens.

PN W0200157274-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI: 2001-488899/53.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT hearts.

PS Claim 15; SEQ ID NO 27753; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for

CC measuring human gene expression in a sample derived from human heart (see

CC ABA21535-ABA41305). The present sequence is a protein encoded by one such

CC probe. The probes may be used for predicting, measuring and displaying

CC gene expression in samples derived from the human heart via microarrays.

CC By measuring gene expression, the probes are useful for predicting,

CC diagnosing, grading, staging, monitoring and prognosing diseases of the

CC human heart and vascular system e.g. cardiovascular disease,

CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 55 AA:

Query Match 2.88; Score 8; DB 22; Length 55;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGIALA 68  
|  
Db 12 LGIGIALA 19

## RESULT 23



```

AAM63528
ID AAM63528 standard; Protein; 55 AA.
XX
AC AAM63528;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35633.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 35633; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 55 AA;
XX
Query Match 2.8%; Score 8; DB 22; Length 55;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 61 LGLGLALA 68
DB 12 LGLGLALA 19
XX
RESULT 24
AAM76342
ID AAM76342 standard; Protein; 55 AA.
XX
AC AAM76342;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36648.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.

```

```

XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 36648; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 55 AA;
XX
Query Match 2.8%; Score 8; DB 22; Length 55;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 61 LGLGLALA 68
DB 12 LGLGLALA 19
XX
RESULT 25
AAM20858
ID AAM20858 standard; Protein; 55 AA.
XX
AC AAM20858;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #1792 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX

```

PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-488901/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human cervical epithelial cells -  
XX  
PS Claim 27; SEQ ID No 25684; 487pp; English.  
XX  
XX The present invention relates to human single exon nucleic acid probes  
CC (SENP: see A110068-A128459). The present sequence is a peptide encoded  
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs  
CC can be used to produce a single exon microarray, which can be used for  
CC measuring human gene expression in a sample derived from human cervical  
CC epithelial cells. By measuring gene expression, the probes are therefore  
CC useful in grading and/or staging of diseases of the cervix, notably  
CC cervical cancer.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).  
XX  
SQ Sequence 55 AA:  
Query Match 2.8%; Score 8; DB 22; Length 55;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 61 LGGLALAA 68  
| | | | | | | | | |  
DB 12 LGGLALAA 19  
RESULT 26  
AAM36451  
ID AAM36451 standard; Protein; 55 AA.  
XX  
AC AAM36451;  
XX  
DT 17-OCT-2001 (first entry)  
XX  
DE Peptide #10488 encoded by probe for measuring placental gene expression.  
XX  
XX Probe; microarray; human; placenta; antenatal diagnosis;  
KM genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200157272-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00663.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2001-48897/53.  
XX  
PT Human genome-derived single exon nucleic acid probes useful for  
PT analyzing gene expression in human placenta -  
XX  
PS Claim 27; SEQ ID No 36720; 654pp; English.

XX  
XX The present invention relates to single exon nucleic acid probes (SENP:  
CC see A113315-A157546). The present sequence is a peptide encoded by one  
CC such probe. The probes are useful for producing a microarray for  
CC predicting, measuring and displaying gene expression in samples derived  
CC from human placenta. The probes are useful for antenatal diagnosis of  
CC human genetic disorders.  
XX  
SQ Sequence 55 AA:  
Query Match 2.8%; Score 8; DB 22; Length 55;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 61 LGGLALAA 68  
| | | | | | | | | |  
DB 12 LGGLALAA 19  
RESULT 27  
ABG45635  
ID ABG45635 standard; Peptide; 55 AA.  
XX  
AC ABG45635;  
XX  
DT 19-AUG-2002 (first entry)  
XX  
DE Human peptide encoded by genome-derived single exon probe SEQ ID 35300.  
XX  
XX Human; single exon probe; asthma; lung cancer; COPD; IID;  
KM chronic obstructive pulmonary disease; interstitial lung disease;  
KM familial idiopathic pulmonary fibrosis; neurofibromatosis;  
KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
KM Heremansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;  
KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
KM primary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
KM primary ciliary dyskinesia; pulmonary hypertension;  
XX hyaline membrane disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200186003-A2.  
XX  
PD 15-NOV-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00665.  
XX  
PR 04-FEB-2000; 2000US-180312P.  
PR 26-MAY-2000; 2000US-207456P.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-234687P.  
PR 27-SEP-2000; 2000US-236359P.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
XX (MOLE-) MOLECULAR DYNAMICS INC.  
PA  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
XX WPI; 2002-114183/15.  
XX  
PT Spatially-addressable set of single exon nucleic acid probes, used to  
PT measure gene expression in human lung samples -  
XX  
XX  
PS Claim 27; SEQ ID No 35300; 634pp; English.  
XX  
XX The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12387 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of  
CC probes; the novel set of probes which hybridise at high stringency to a

CC nucleic acid expressed in the human lung; measuring gene expression in a  
CC sample derived from human lung, comprising (a) contacting the array with  
CC a collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of  
CC the array; identifying exons in a eukaryotic genome, comprising  
CC (a) algorithmically predicting at least one exon from genomic sequences  
CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray, assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridisation to a single exon  
CC microarrays having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene  
CC expression analysis, and for identifying exons in a gene, particularly  
CC using human lung derived mRNA and for the study of lung diseases  
CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary  
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic  
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension  
CC and hyaline membrane disease. The present sequence is a peptide/protein  
CC encoded by a single exon probe of the invention.  
CC Note: The sequence data for this patent did not form part  
CC of the printed specification, but was obtained in electronic  
CC format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 55 AA:

Query Match 2.8%; Score 8; DB 23; Length 55;  
Best Local Similarity 100.0%; Pred. No. 13;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGLGLALA 68  
          |||||  
DB 12 LGLGLALA 19

RESULT 28

ABBA43981  
ID ABB43981 standard; Peptide: 65 AA.

XX ABB43981;

XX 04-FEB-2002 (first entry)

DE Peptide #11487 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.

OS WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00669.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

DR Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human fetal liver -

XX Claim 27; SEQ ID NO 36616; 639bp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for  
CC measuring human gene expression in a sample derived from human foetal  
CC liver. The single exon nucleic acid probes may be used for predicting,  
CC measuring and displaying gene expression in samples derived from human  
CC fetal liver. The present sequence is a peptide encoded by a single exon  
CC nucleic acid probe of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 65 AA:

Query Match 2.8%; Score 8; DB 22; Length 65;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGLGLALA 68  
          |||||  
DB 22 LGLGLALA 29

RESULT 29

ABB26890  
ID ABB26890 standard; Protein: 65 AA.

XX ABB26890;

XX 23-JAN-2002 (first entry)

DE Protein #8889 encoded by probe for measuring heart cell gene expression.

XX Human; gene expression; heart; microarray; vascular system;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease.

XX Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00666.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

DR Single exon nucleic acid probes for analyzing gene expression in human

PT hearts -

XX Claim 15; SEQ ID NO 28660; 530bp; English.

CC The present invention relates to single exon nucleic acid probes for  
CC measuring human gene expression in a sample derived from human heart (see  
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such  
CC probe. The probes may be used for predicting, measuring and displaying  
CC gene expression in samples derived from the human heart via microarrays.  
CC By measuring gene expression, the probes are useful for predicting,  
CC diagnosing, grading, staging, monitoring and prognosing diseases of the  
CC human heart and vascular system e.g. cardiovascular disease,  
CC hypertension, cardiac arrhythmias and congenital heart disease.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 65 AA;

Query Match 2.88; Score 8; DB 22; Length 65;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGTGLALA 68  
|||  
Db 22 LGTGLALA 29

RESULT 30

AAAM64984 standard; Protein: 65 AA.

XX ID AAM64984;

XX AC AAM64984;

XX DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 37089.  
XX Human: brain expressed exon; gene expression analysis; probe;  
KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
XX epilepsy; cancer.

XX OS Homo sapiens.

XX PN MO200157275-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00667.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483446/52.

XX PT Single exon nucleic acid probes for analyzing gene expression in human

XX PS brains -

XX Example 4; SEQ ID NO: 37089; 650bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX brain. They can be used to measure gene expression in brain cell samples,  
XX which may enable the diagnosis and improved treatment of nervous system  
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
XX epilepsy and cancers. The present sequence is a protein encoded by one of  
XX the probes of the invention.

SQ Sequence 65 AA;

Query Match 2.88; Score 8; DB 22; Length 65;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGTGLALA 68  
|||  
Db 22 LGTGLALA 29

RESULT 31

AAAM77706 standard; Protein: 65 AA.

XX ID AAM77706;

XX AC AAM77706;

XX DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 38012.  
XX Human: bone marrow expressed exon; gene expression analysis; probe;  
KM microarray; cancer; leukemia; lymphoma; myeloma.

XX OS Homo sapiens.

XX PN MO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00668.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488900/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PS analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 38012; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
XX bone marrow. They can be used to measure gene expression in bone marrow  
XX samples, which may enable the improved diagnosis and treatment of cancers  
XX such as lymphoma, leukemia and myeloma. The present sequence is a  
XX protein encoded by one of the probes of the invention.

XX SQ Sequence 65 AA;  
Query Match 2.88; Score 8; DB 22; Length 65;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGTGLALA 68  
|||  
Db 22 LGTGLALA 29

RESULT 32

AAAM21621 standard; Protein: 65 AA.

XX ID AAM21621

AC AAM21621;  
 XX  
 DN 12-OCT-2001 (first entry)  
 XX  
 DE Peptide #8055 encoded by probe for measuring cervical gene expression.  
 XX  
 KW Probe; human; microarray; gene expression; cervical epithelial cell;  
 KW cervical cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157278-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00670.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488901/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human cervical epithelial cells -  
 XX  
 PS Claim 27; SEQ ID No 26447; 487bp; English.  
 XX  
 CC The present invention relates to human single exon nucleic acid probes  
 CC (SNP; see A110068-A128459). The present sequence is a peptide encoded  
 CC by one such probe. The SNPs are derived from human HeLa cells. The SNPs  
 CC can be used to produce a single exon microarray, which can be used for  
 CC measuring human gene expression in a sample derived from human cervical  
 CC epithelial cells. By measuring gene expression, the probes are therefore  
 CC useful in grading and/or staging of diseases of the cervix, notably  
 CC cervical cancer.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SQ Sequence 65 AA;  
 Query Match 2.8%; Score 8; DB 22; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 61 LGLGIALA 68  
 |||||  
 DB 22 LGLGIALA 29  
 RESULT 33  
 AAM37923  
 ID AAM37923 standard; Protein; 65 AA.  
 XX  
 AC AAM37923;  
 XX  
 DN 17-OCT-2001 (first entry)  
 XX  
 DE Peptide #11960 encoded by probe for measuring placental gene expression.  
 XX  
 KW Probe; microarray; human; placenta; antenatal diagnosis;  
 KW genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX

XX  
 PN WO200157272-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00663.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000GB-0024263.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488897/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -  
 XX  
 PS Claim 27; SEQ ID No 38192; 654bp; English.  
 XX  
 CC The present invention relates to single exon nucleic acid probes (SNP;  
 CC see A113315-A157546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.  
 XX  
 SQ Sequence 65 AA;  
 Query Match 2.8%; Score 8; DB 22; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 61 LGLGIALA 68  
 |||||  
 DB 22 LGLGIALA 29  
 RESULT 34  
 ABG46737  
 ID ABG46737 standard; Peptide; 65 AA.  
 XX  
 AC ABG46737;  
 XX  
 DN 19-AUG-2002 (first entry)  
 XX  
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 36402.  
 XX  
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200186003-A2.  
 XX  
 PD 15-NOV-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00665.  
 XX  
 PR 04-FEB-2000; 2000US-180312P.  
 XX

PR 26-MAY-2000; 2000US-207456P.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-234687P.  
 PR 27-SEP-2000; 2000US-236359P.  
 PR 04-OCT-2000; 2000US-0024263.  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI; 2002-114183/15.  
 DR Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples -  
 XX  
 PS Claim 27; SEQ ID No 36402; 634pp; English.  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of  
 CC probes; the novel set of probes which hybridise at high stringency to a  
 CC nucleic acid expressed in the human lung; measuring gene expression in a  
 CC sample derived from human lung, comprising (a) contacting the array with  
 CC a collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of  
 CC the array; identifying exons in a eukaryotic genome, comprising  
 CC (a) algorithmically predicting at least one exon from genomic sequences  
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene  
 CC expression analysis, and for identifying exons in a gene, particularly  
 CC using human lung derived mRNA and for the study of lung diseases  
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease  
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary  
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,  
 CC Niemann-Pick disease, Hermansky-Rudlak syndrome, sarcoidosis, pulmonary  
 CC haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,  
 CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic  
 CC and hyaline membrane disease. The present sequence is a peptide/protein  
 CC encoded by a single exon probe of the invention  
 CC Note: The sequence data for this patent did not form part  
 CC of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SO Sequence 65 AA;  
 Query Match 2.8%; Score 8; DB 23; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 61 LGIGLALA 68  
 DB 22 LGIGLALA 29  
 RESULT 35  
 ABG03723  
 ID ABG03723 standard; Protein; 69 AA.

XX  
 AC ABG03723;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE Novel human diagnostic protein #3714.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 PF 30-MAR-2001; 2001MO-US08631.  
 XX  
 PR 31-MAR-2000; 2000US-0540217.  
 PR 23-AUG-2000; 2000US-0649167.  
 XX  
 PA (HYSE-) HYSER INC.  
 XX  
 PI Dymnac RT, Liu C, Tang YT;  
 XX  
 DR WPI; 2001-639362/73.  
 DR N-PSDB; AAS67910.  
 PS Claim 20; SEQ ID No 34082; 103pp; English.  
 CC The invention relates to isolated polynucleotide (I) and  
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
 CC and gene mapping, and in recombinant production of (II). The  
 CC polynucleotides are also used in diagnostics as expressed sequence tags  
 CC for identifying expressed genes. (I) is useful in gene therapy techniques  
 CC to restore normal activity of (II) or to treat disease states involving  
 CC (II). (II) is useful for generating antibodies against it, detecting or  
 CC quantitating a polypeptide in tissue, as molecular weight markers and as  
 CC a food supplement. (II) and its binding partners are useful in medical  
 CC imaging of sites expressing (II). (I) and (II) are useful for treating  
 CC disorders involving aberrant protein expression or biological activity.  
 CC The polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. ABG00010-ABG30377 represent novel human  
 CC diagnostic amino acid sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 XX  
 SO Sequence 69 AA;  
 Query Match 2.8%; Score 8; DB 22; Length 69;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 61 LGIGLALA 68  
 DB 16 LGIGLALA 23  
 RESULT 36  
 ABG03663  
 ID ABG03663 standard; Protein; 71 AA.  
 AC ABG03663;

XX 13-FEB-2002 (first entry)  
DT  
XX  
DE Novel human diagnostic protein #3654.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PE 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI: 2001-639362/73.  
DR N-PSDB: AAS67850.  
XX  
XX  
PS Claim 20; SEQ ID No 34022; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG0010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 71 AA:  
Query Match 2.8%; Score 8; DB 22; Length 71;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 61 LGIGIALA 68  
DB 16 LGIGIALA 23  
RESULT 37  
AAU61282  
ID AAU61282 standard; Protein: 84 AA.  
AC AAU61282;  
XX  
XX 27-FEB-2002 (first entry)  
DT

XX Propionibacterium acnes immunogenic protein #22178.  
DE  
XX  
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;  
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;  
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;  
KW dermatological; osteopathic; neuroprotectant.  
XX  
XX Propionibacterium acnes.  
XX  
PN WO200181581-A2.  
XX  
PD 01-NOV-2001.  
XX  
PE 20-APR-2001; 2001WO-US12865.  
XX  
PR 21-APR-2000; 2000US-199047P.  
PR 02-JUN-2000; 2000US-208841P.  
PR 07-JUL-2000; 2000US-216747P.  
XX  
XX (CORI-) CORIXA CORP.  
XX  
XX Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;  
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;  
XX  
DR WPI: 2001-616774/71.  
DR N-PSDB: AAS59615.  
XX  
XX  
PS Example 1; SEQ ID No 22477; 1069pp; English.  
XX  
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 84 AA:  
Query Match 2.8%; Score 8; DB 22; Length 84;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 259 LRIRLTPW 266  
DB 34 LRIRLTPW 41  
RESULT 38  
ABG20259  
ID ABG20259 standard; Protein: 110 AA.  
AC ABG20259;  
XX  
XX 18-FEB-2002 (first entry)  
DT





PD 27-SEP-2001.  
XX  
XX 23-MAR-2001; 2001WO-US09231.  
XX  
XX 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX  
XX (PERE ) PE CORP NY.  
PA  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
XX WPI; 2001-656860/75.  
DR N-PSDB; ABL11815.  
XX  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PI interactions -  
XX  
XX  
PS Disclosure; SEQ ID NO 29928; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (AB57737-AB872072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 184 AA;  
  
Query Match 2.8%; Score 8; DB 22; Length 184;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 72 LLLAVVSL 79  
DB 16 LLLAVVSL 23  
  
RESULT 41  
ABP28041  
ID ABP28041 standard; Protein; 190 AA.  
XX  
XX ABP28041;  
AC  
XX  
XX 02-JUL-2002 (first entry)  
DT  
XX  
XX Streptococcus polypeptide SEQ ID NO 5258.  
DE  
XX  
XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;  
KW group A streptococcus; Streptococcus pyogenes; antibacterial;  
KW antiinflammatory; Infection; vaccine; meningitis; gene therapy.  
XX  
XX Streptococcus agalactiae.  
OS  
XX  
XX WO200234771-A2.  
PN  
XX  
XX 02-MAY-2002.  
PD  
XX  
XX 29-OCT-2001; 2001WO-GB04789.  
PF  
XX  
XX 27-OCT-2000; 2000GB-0026333.  
PR 24-NOV-2000; 2000GB-0028727.  
PR 07-MAR-2001; 2001GB-0005640.  
XX  
XX (CHIR-) CHIRON SPA.  
PA (GENO-) INST GENOMIC RES.  
PI Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;

PI Tettelin H;  
XX  
XX WPI; 2002-352536/38.  
DR N-PSDB; ABN68672.  
XX  
XX New Streptococcus protein for the treatment or prevention of infection  
PT or disease caused by Streptococcus bacteria, such as meningitis, and  
PI for detecting a compound that binds to the protein -  
XX  
XX  
PS Claim 1; Page 3689; 4525pp; English.  
XX  
XX The invention relates to a protein (ABP25413-ABP30895) from group B  
CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), ABN6804-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins.  
XX  
XX  
SQ Sequence 190 AA;  
  
Query Match 2.8%; Score 8; DB 23; Length 190;  
Best Local Similarity 100.0%; Pred. No. 44;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 258 SLRIRTLR 265  
DB 50 SLRIRTLR 57  
  
RESULT 42  
AAM23684  
ID AAM23684 standard; Protein; 198 AA.  
XX  
XX AAM23684;  
AC  
XX  
XX 12-OCT-2001 (first entry)  
DT  
XX  
XX Human EST encoded protein SEQ ID NO: 1209.  
DE  
XX  
XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
KW diagnostics; forensic test; gene mapping; genetic disorder;  
KW biodiversity; gene therapy; nutrition.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200154477-A2.  
PN  
XX  
XX 02-AUG-2001.  
PD  
XX  
XX 25-JAN-2001; 2001WO-US02687.  
PF  
XX  
XX 25-JAN-2000; 2000US-0491404.  
PR 17-JUL-2000; 2000US-0617746.  
PR 03-AUG-2000; 2000US-0631451.  
PR 15-SEP-2000; 2000US-0663870.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX  
XX Tang YT, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Drmanac RA, Zhang J, Werhman T;

DR WPI: 2001-476164/51.  
DR N-PSDB: AAB98343.  
XX  
PT Isolated polypeptide for treatment of diseases, diagnostics, raising  
PT antibodies and research use -  
XX  
PS Claim 20; Page 873; 1275pp; English.  
XX  
CC The present invention provides the protein and coding sequences of novel  
CC proteins from a variety of organisms, including human, dog, cat, horse,  
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
CC from the organism of interest. They can be used in diagnostics,  
CC forensics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a  
CC protein of the invention.  
XX  
SQ Sequence 198 AA:  
Query Match 2.8%; Score 8; DB 22; Length 198;  
Best Local Similarity 100.0%; Pred. No. 45;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 54 ALVPLAL 61  
|||||||  
DB 2 ALVPLAL 9  
RESULT 43  
ABP41674  
ID ABP41674 standard; Protein: 222 AA.  
XX  
AC ABP41674;  
XX  
DT 22-AUG-2002 (first entry)  
XX  
DE Human ovarian antigen HPAMG11, SEQ ID NO:2806.  
XX  
KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
KW PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;  
KW inflammatory condition; immune disorder; blood disorder;  
KW cardiovascular disorder; respiratory disorder; neurological disorder;  
KW gastrointestinal disorder; urinary system disorder; drug screening;  
KW gene therapy; chromosome mapping; forensic analysis;  
KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
KW antiinflammatory; gynaecological; reproductive.  
XX  
OS Homo sapiens.  
XX  
PN WO200200677-A1.  
XX  
PD 03-JAN-2002.  
XX  
PF 07-JUN-2001; 2001WO-US18569.  
XX  
PR 07-JUN-2000; 2000US-209467P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Blrse CE, Rosen CA;  
XX  
DR WPI: 2002-147878/19.  
DR N-PSDB: ABQ54751.  
XX  
PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
PT useful in the prevention, treatment and diagnosis of cancer (e.g.  
PT ovarian cancer), immune disorders, cardiovascular disorders and  
PT neurological diseases -  
XX  
PS claim 11; SEQ ID NO 2806; 2922pp; English.  
XX

CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also  
CC encompasses polypeptides 90% identical and polynucleotides 95% identical  
CC to the sequences of the invention. The invention additionally relates to  
CC recombinant vectors and host cells comprising human ovarian antigen  
CC polynucleotides, antibodies against human ovarian antigens, and the use  
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
CC treating, prognosing or preventing various ovary and/or breast-related  
CC disorders. Such conditions include ovarian cancer and breast cancer, and  
CC metastatic tumours of ovarian or breast origin, reproductive system  
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine  
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
CC vaginitis), immune disorders (e.g., congenital and acquired  
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
CC respiratory disorders, neurological disorders, gastrointestinal disorders  
CC and urinary system disorders. Ovarian antigen polypeptides and  
CC polynucleotides may also be used in screening for compounds which  
CC modulate ovarian antigen expression or activity. The polynucleotides may  
CC further be used for gene therapy, chromosome mapping, in the  
CC identification of individuals and in forensic analysis, and the  
CC polypeptides may be used as food additives or to prepare antibodies  
CC useful in disease diagnosis, drug targeting and phenotyping. The present  
CC sequence represents a human ovarian antigen of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 222 AA:

Query Match 2.8%; Score 8; DB 23; Length 222;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGIGLAL 67  
|||||||  
DB 190 ALGIGLAL 197

RESULT 44  
ABG16279  
ID ABG16279 standard; Protein: 286 AA.  
XX  
AC ABG16279;  
XX  
DT 18-FEB-2002 (first entry)  
XX

DE Novel human diagnostic protein #16270.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX

PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
XX  
PR 23-AUG-2000; 2000US-0649167.  
XX

PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI: 2001-639362/73.  
DR N-PSDB: AAS80466.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity  
XX  
PS Claim 20; SEQ ID No 46638; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 286 AA;  
Query Match 2.8%; Score 8; DB 22; Length 286;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 251 LALRPGSS 258  
|||||  
Db 69 LALRPGSS 76

RESULT 45  
ABG01186  
ID ABG01186 standard; Protein; 307 AA.  
XX  
AC ABG01186;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #1177.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI: 2001-639362/73.  
DR N-PSDB; AAS65373.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess

PT biodiversity  
XX  
PS Claim 20; SEQ ID No 31545; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 307 AA;  
Query Match 2.8%; Score 8; DB 22; Length 307;  
Best Local Similarity 100.0%; Pred. No. 69;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 77 VSLGSRAS 84  
|||||  
Db 3 VSLGSRAS 10

Search completed: May 8, 2003, 06:55:33  
Job time : 65 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.5  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 8, 2003, 06:54:40 ; Search time: 46 Seconds

(without alignments)  
593.526 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 284  
Sequence: 1 MSLLDFEISARLPLPRSLG.....PMAHLKAPFLTYGFLQVH 284

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: PIR\_73.\*

1: PIR1.\*  
2: PIR2.\*  
3: PIR3.\*  
4: PIR4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	3.5	111	2 A85866	hypothetical prote
2	10	3.2	733	2 S78376	photosystem I P700
3	8	2.8	58	2 A58208	protamine I-1 - pa
4	8	2.8	143	2 G84168	hypothetical prote
5	8	2.8	197	2 E72374	hypothetical prote
6	8	2.8	220	2 AG3547	biclyclomycin resis
7	8	2.8	278	2 D83080	hypothetical prote
8	8	2.8	339	2 C71132	hypothetical prote
9	8	2.8	372	2 H70813	formate dehydrogen
10	8	2.8	379	2 E64300	conserved hypotnet
11	8	2.8	381	2 AH3041	hypothetical prote
12	8	2.8	387	2 D84885	hypothetical prote
13	8	2.8	397	2 D98244	hypothetical prote
14	8	2.8	422	2 A60503	sperm-binding glyco
15	8	2.8	443	2 T17220	hypothetical prote
16	8	2.8	465	2 AC0347	probable membrane
17	8	2.8	471	2 A75267	probable transpor
18	8	2.8	576	2 E64186	gamma-glutamyltran
19	8	2.8	586	2 A41125	protamine II-3 - p
20	7	2.5	45	2 D58208	protamine I - Sagu
21	7	2.5	50	2 S22582	protamine I - North
22	7	2.5	58	2 S34045	hypothetical prote
23	7	2.5	86	2 F87604	protein ZC334.3 [1
24	7	2.5	102	2 F87993	ig heavy chain V r
25	7	2.5	115	2 PH1560	conserved hypotnet
26	7	2.5	115	2 H83201	hypothetical prote
27	7	2.5	118	1 IECSB	Orf51 [bacterioph
28	7	2.5	118	2 AE1753	conserved hypotnet
29	7	2.5	123	2 AH2707	conserved hypotnet

30	7	2.5	125	2 T27519	hypothetical prote
31	7	2.5	131	2 I52290	interleukin-13 - r
32	7	2.5	131	2 E30552	T-cell activation
33	7	2.5	146	2 T37116	probable transpos
34	7	2.5	147	2 A71217	hypothetical prote
35	7	2.5	150	2 T08734	hypothetical prote
36	7	2.5	157	2 S31078	seed allergen RA5
37	7	2.5	157	2 T02664	allergen - rice
38	7	2.5	157	2 A75567	conserved hypotnet
39	7	2.5	157	2 E75350	hypothetical prote
40	7	2.5	160	2 S59925	allergen RA5B prec
41	7	2.5	161	2 DNEC17	outer membrane pro
42	7	2.5	161	2 D90651	histone-like prote
43	7	2.5	161	2 D85502	hypothetical prote
44	7	2.5	162	2 T24937	hypothetical prote
45	7	2.5	162	2 T31173	hypothetical prote

#### ALIGNMENTS

RESULT 1  
A85866  
hypothetical protein Z3516 [imported] - Escherichia coli (strain O157:H7, substrain  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: A85866  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Ma  
iller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamotis, K.; Apod  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: A85866  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-111 <STO>  
A:Cross-references: GB:AE005174; NID:q12516604; PIDN:AAG57389.1; GSPDB:GN00145; UWGP  
C:Experimental source: strain O157:H7, substrain EDL933  
A:Genetics:  
A:Gene: Z3516

Query Match 3.5%; Score 10; DB 2; Length 111;  
Best Local Similarity 100.0%; Pred. No. 0.073;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 IGLAIACTGL 72  
DB 40 IGLAIACTGL 49  
|||||  
|||

RESULT 2  
S78376  
Photosystem I P700 apoprotein A2 - Odontella sinensis chloroplast  
C:Species: chloroplast Odontella sinensis  
C:Date: 17-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 20-Jun-2000  
C:Accession: S78376  
R:Kowallik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Panolic, P.; Freiler, U.  
Plant Mol. Biol. Rep. 13, 336-342, 1995  
A:Title: The Chloroplast Genome of a Chlorophyll a+c- containing Alga, Odontella sin  
A:Reference number: S78238  
A:Accession: S78376  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-733 <KOM>  
A:Cross-references: EMBL:Z67753; NID:q1185127; PIDN:CAA91749.1; PID:q1185266  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 199  
C:Genetics:  
A:Gene: psab  
A:Genome: chloroplast  
C:Superfamily: photosystem I P700 apoprotein  
C:Keywords: chloroplast; electron transfer; membrane protein; membrane-associated co

Query Match 3.2%; Score 9; DB 2; Length 733;

Best Local Similarity 100.0%; Pred. No. 3.3;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 63 LGIALLACIG 71  
|||||  
Db 333 LGIALLACIG 341

RESULT 3  
A58208  
protamine I-1 - painted turtle  
C:Species: Chrysemys picta (painted turtle)  
C:Date: 08-Nov-1996 #sequence\_revision 08-Nov-1996 #text\_change 07-May-1999  
C:Accession: A58208  
R:Hunt, J.G.; Kasinsky, H.E.; Elsey, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp, D.  
J. Biol. Chem. 271, 23547-23557, 1996  
A:Title: Protamines of reptiles.  
A:Reference number: A58208; MID:96394458; PMID:8798564  
A:Accession: A58208  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-58 <HUN>  
C:Superfamily: sperm histone

Query Match 2.8%; Score 8; DB 2; Length 58;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 42 QRRRGRRG 49  
|||||  
Db 35 QRRRGRRG 42

RESULT 4  
G84168  
hypothetical protein Vng0080h [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: G84168  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.  
; Leitauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabcic  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MID:20504483; PMID:11016950  
A:Accession: G84168  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-143 <STO>  
A:Cross-references: GB:AE004437; NID:g10579733; PIDN:AA618715.1; GSPDB:GN00138  
C:Genetics:  
A:Gene: VNG0080H

Query Match 2.8%; Score 8; DB 2; Length 143;  
Best Local Similarity 100.0%; Pred. No. 7.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 61 LGIGLALA 68  
|||||  
Db 55 LGIGLALA 62

RESULT 5  
E72374  
hypothetical protein - Thermotoga maritima (strain MSB8)  
C:Species: Thermotoga maritima  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: E72374  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;  
C.M.  
Nature 399, 323-329, 1999

A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome  
A:Reference number: A72200; MID:99287316; PMID:10360571  
A:Accession: E72374  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-197 <ARN>  
A:Cross-references: GB:AE001724; GB:AE000512; NID:g4980966; PIDN:AA035553.1; PID:g498  
C:Genetics:  
A:Gene: TM0469

Query Match 2.8%; Score 8; DB 2; Length 197;  
Best Local Similarity 100.0%; Pred. No. 9.9;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 SLDDFEIS 9  
|||||  
Db 135 SLDDFEIS 142

RESULT 6  
AG3547  
bicyclomycin resistance protein [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: AG3547  
R:Belvecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Muter, C.; Los, T.; Ivanov  
; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melit  
A:Reference number: AD252; PMID:1175668  
A:Accession: AG3547  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-220 <KUR>  
A:Cross-references: GB:AE008918; PIDN:AA153546.1; PID:g17984455; GSPDB:GN00191  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI10304  
A:Map position: II

Query Match 2.8%; Score 8; DB 2; Length 220;  
Best Local Similarity 100.0%; Pred. No. 11;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 54 ALLVPLAL 61  
|||||  
Db 132 ALLVPLAL 139

RESULT 7  
D83080  
hypothetical protein PA4521 [imported] - Pseudomonas aeruginosa (strain PA01)  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: D83080  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; L  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
A:Reference number: AB2950; MID:20437337; PMID:10984043  
A:Accession: D83080  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-278 <STO>  
A:Cross-references: GB:AE004866; GB:AE004091; NID:g9950760; PIDN:AA07909.1; GSPDB:GN  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA4521

Query Match 2.8%; Score 8; DB 2; Length 278;  
Best Local Similarity 100.0%; Pred. No. 13;

Matches 8: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 LGLLAVV 77  
 DB 47 LGLLAVV 54

RESULT 8  
 C71132  
 hypothetical protein PH0824 - *Pyrococcus horikoshii*  
 C:Species: *Pyrococcus horikoshii*  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
 C:Accession: C71132  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Seki, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic *A. ferrireducens* strain  
 A:Reference number: A71000; MUID:98344137; PMID:9679194  
 A:Accession: C71132  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-339 <KAP>  
 A:Cross-references: GB:AP000003; NID:93236130; PIDN:BAA29917.1; PID:93257234  
 A:Experimental source: strain 073  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C:Genetics:  
 A:Gene: PH0824  
 C:Superfamily: conserved hypothetical protein MTR900

Query Match 2.8%; Score 8; DB 2; Length 339;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 LGLLAVV 77  
 DB 111 LGLLAVV 118

RESULT 9  
 H70813  
 probable cysteine synthase - *Mycobacterium tuberculosis* (strain H37Rv)  
 C:Species: *Mycobacterium tuberculosis*  
 C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
 C:Accession: H70813  
 R:Conor, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome  
 A:Reference number: A70500; MUID:98295987; PMID:9634230  
 A:Accession: H70813  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-312 <COB>  
 A:Cross-references: GB:AL022004; GB:AL124345; NID:93261550; PIDN:CAAL7654.1; PID:9291690  
 A:Experimental source: strain H37Rv  
 C:Genetics:  
 A:Gene: cysM3  
 C:Superfamily: threonine dehydratase

Query Match 2.8%; Score 8; DB 2; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LGIGLALA 68  
 DB 96 LGIGLALA 103

RESULT 10  
 E64300  
 formate dehydrogenase (EC 1.2.1.2) beta chain - *Methanococcus jannaschii*

C:Species: *Methanococcus jannaschii*  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
 C:Accession: E64300  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blais, J.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, I.; Roun, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996  
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woes, A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*  
 A:Reference number: A64300; MUID:96337999; PMID:8688087  
 A:Accession: E64300  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-379 <BUT>  
 A:Cross-references: GB:U67459; GB:L77117; NID:92826236; PIDN:AAB97986.1; PID:9159081  
 C:Genetics:  
 A:Map position: REV7250-6111  
 C:Superfamily: formate dehydrogenase chain B; ferredoxin 2[4Fe-4S] homology  
 C:Keywords: oxidoreductase  
 F:273-348/Domain: ferredoxin 2[4Fe-4S] homology <FE>

Query Match 2.8%; Score 8; DB 2; Length 379;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 217 LVDGVVAL 224  
 DB 35 LVDGVVAL 42

RESULT 11  
 AH3041  
 conserved hypothetical protein Atu3948 [imported] - *Agrobacterium tumefaciens* (strain AH3041)  
 C:Species: *Agrobacterium tumefaciens*  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
 C:Accession: AH3041  
 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; McRae, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; MCC, Karp, P.; Romero, P.; Zhang, S. Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kester, E.W.  
 A:Title: The genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
 A:Reference number: AB2577; PMID:11743193  
 A:Accession: AH3041  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-381 <KUR>  
 A:Cross-references: GB:AE008689; PIDN:AAL44750.1; PID:917742385; GSPDB:GN00187  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atu3948  
 A:Map position: linear chromosome

Query Match 2.8%; Score 8; DB 2; Length 381;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 251 LALRPGSS 258  
 DB 325 LALRPGSS 332

RESULT 12  
 D84885  
 hypothetical protein At2g45000 [imported] - *Arabidopsis thaliana*  
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
 C:Accession: D84885  
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C. M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: D84885  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-387 <STO>  
A:Cross-references: GB:AE002093; NID:94895250; PIDN:AAD32835.1; GSPDB:GN00139  
A:Gene: At2g45000  
A:Map position: 2

Query Match 2.8%; Score 8; DB 2; Length 387;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 93 ELVAEED 100  
DB 370 ELVAEED 377

## RESULT 13

D98244

hypothetical protein AGR\_L1808 [imported] - *Agrobacterium tumefaciens* (strain C58, Cere  
C:Species: *Agrobacterium tumefaciens*  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002  
C:Accession: D98244  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouello, B.; Goldman,  
A.; Liu, F.; Mollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tum*  
A:Reference number: A97359; PMID:11743194  
A:Accession: D98244  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-397 <KUR>  
A:Cross-references: GB:AE007870; PIDN:AAK89478.1; PID:q15159347; GSPDB:GN00170  
C:Genetics:  
A:Gene: AGR\_L1808  
A:Map position: linear chromosome

Query Match 2.8%; Score 8; DB 2; Length 397;  
Best Local Similarity 100.0%; Pred. No. 18;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 251 LALRGS 258  
DB 341 LALRGS 348

## RESULT 14

A60503

sperm-binding glycoprotein ZP3 precursor - golden hamster  
N:Alternate names: sperm receptor; zona pellucida glycoprotein ZP3  
C:Species: *Mesocricetus auratus* (golden hamster)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A60503  
R:Kunlich, R.A.; Ruiz-Seller, B.; Massarman, P.M.  
Dev. Biol. 142, 414-421, 1990  
A:Title: Genomic organization and polypeptide primary structure of zona pellucida glycop  
A:Reference number: A60503; MUID:91078540; PMID:2257975  
A:Accession: A60503  
A:Molecule type: DNA  
A:Residues: 1-422 <KIN>  
A:Cross-references: GB:M63629  
A:Note: the authors translated the codon CAA for residue 251 as Glu, and AGG for residue  
C:Comment: This sulfated glycoprotein in the zona pellucida of the oocyte is a receptor  
C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology  
C:Keywords: glycoprotein; oocyte  
F:45-300/Domain: ZP domain homology <ZPH>

Query Match 2.8%; Score 8; DB 1; Length 422;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LALGLA 66  
DB 386 LALGLA 393

## RESULT 15

T17220

hypothetical protein DKFZp566001.1 - human  
C:Species: *Homo sapiens* (man)  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 02-Jun-2000  
C:Accession: T17220  
R:Bium, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, September 1999  
A:Reference number: Z18725  
A:Accession: T17220  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-443 <BLU>  
A:Cross-references: EMBL:AL117414  
A:Experimental source: fetal kidney; clone DKFZp566001  
C:Genetics:  
A:Note: DKFZp566001.1  
C:Superfamily: gamma-glutamyltransferase

Query Match 2.8%; Score 8; DB 2; Length 443;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGGLALA 68  
DB 14 LGGLALA 21

## RESULT 16

AC0347

probable membrane protein yegB [imported] - *Yersinia pestis* (strain CO92)  
C:Species: *Yersinia pestis*  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 17-May-2002  
C:Accession: AC0347  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittall, R.W.; Holden, M.T.G.; Prentice, M.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.  
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barral  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AC0347  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-465 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC92102.1; PID:q15980820; GSPDB:GN00175  
C:Genetics:  
A:Gene: yegB  
C:Superfamily: multidrug-efflux transporter

Query Match 2.8%; Score 8; DB 2; Length 465;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLAIVSL 79  
DB 335 LLAIVSL 342

## RESULT 17

A75267

probable transport protein - *Deinococcus radiodurans* (strain R1)  
C:Species: *Deinococcus radiodurans*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
C:Accession: A75267  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.  
M.; Shen, M.; Vamathevan, J.C.; Lam, P.; McDonald, L.; Uitterlind, T.; Zalewski, C.;  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.



Science 286, 1571-1577, 1999.

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: A75267

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-471 <MAN>

A:Cross-references: GB:AE002079; GB:AE000513; NID:g6460315; PIDN:AAFI2043.1; PID:g646032

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR2502

A:Map position: 1

Query Match 2.8%; Score 8; DB 2; Length 471;

Best Local Similarity 100.0%; Pred. No. 21;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGIAL 67

DB 366 ALGIAL 373

RESULT 18

E64186

probable ATP-binding transport protein H1156 - *Haemophilus influenzae* (strain Rd KW20)

C:Species: *Haemophilus influenzae*

C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 02-Feb-2001

C:Accession: E64186

R:Flatschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.

R:Goatley, J.D.; Scott, J.; Shiley, R.; Liu, L.I.; Glodok, A.; Kelley, J.M.; Weidman, J.

R:D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhmann, J.L.; Geoghegan, N.S.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.

A:Reference number: A64000; MUID:9550630; PMID:7542800

A:Accession: E64186

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-576 <TIGR>

A:Cross-references: GB:U3795; GB:LA2023; NID:g1574708; PIDN:AAC2281.1; PID:g1574713;

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

C:Keywords: ATP; nucleotide binding; P-loop

F:355-550/Domain: ATP-binding cassette homology <ABC>

F:372-379/Region: nucleotide-binding motif A (P-loop)

Query Match

Best Local Similarity 2.8%; Score 8; DB 2; Length 576;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 58 PLALGAL 65

DB 159 PLALGAL 166

RESULT 19

AA1125

gamma-glutamyltransferase (EC 2.3.2.2) related protein - human

N:Alcarnate names: gamma-glutamyltransferase-like activity 1; GGT-REL

C:Species: *Homo sapiens* (man)

C>Date: 27-Mar-1992 #sequence\_revision 27-Mar-1992 #text\_change 18-Jun-1999

C:Accession: AA1125

R:Helsterkamp, N.; Rajpert-De Meyts, E.; Uribe, L.; Forman, H.J.; Groffen, J.

Proc. Natl. Acad. Sci. U.S.A. 88, 6303-6307, 1991

A:Title: Identification of a human gamma-glutamyl cleaving enzyme related to, but distinct

A:Reference number: AA1125; MUID:91296809; PMID:1676842

A:Accession: AA1125

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-586 <HEI>

A:Cross-references: GB:M64099; NID:g183141; PIDN:AA58503.1; PID:g183142

C:Genetics:

A:Gene: GDB:GGTAL1; GGT-REL

A:Cross-references: GDB:134033

C:Superfamily: gamma-glutamyltransferase

C:Keywords: aminoacyltransferase; glycoprotein; transmembrane protein

Query Match 2.8%; Score 8; DB 2; Length 586;

Best Local Similarity 100.0%; Pred. No. 25;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68

DB 14 LGIGLALA 21

RESULT 20

D58208

protamine II-3 - painted turtle

C:Species: *Chrysemys picta* (painted turtle)

C>Date: 08-Nov-1996 #sequence\_revision 08-Nov-1996 #text\_change 07-May-1999

C:Accession: D58208

R:Hunt, J.G.; Kasinsky, H.E.; Elsey, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp

J. Biol. Chem. 271, 23547-23557, 1996

A:Title: Protamines of reptiles.

A:Reference number: A58208; MUID:96394458; PMID:8798564

A:Accession: D58208

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-45 <HUN>

C:Superfamily: sperm histone

Query Match 2.5%; Score 7; DB 2; Length 45;

Best Local Similarity 100.0%; Pred. No. 26;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRG 49

DB 24 RRRGRG 30

RESULT 21

S22582

protamine 1 - *Saguinus imperator*

C:Species: *Saguinus imperator*

C>Date: 22-Nov-1993 #sequence\_revision 12-Apr-1996 #text\_change 21-Jul-2000

C:Accession: S22582

R:Queralto, R.; Oliva, R.

Nucleic Acids Res. 19, 5786, 1991

A:Title: Protamine 1 gene sequence from the primate *Saguinus imperator* isolated with

A:Reference number: S22582; MUID:92051332; PMID:1840669

A:Accession: S22582

A:Molecule type: DNA

A:Residues: 1-50 <QUE>

A:Cross-references: EMBL:X61678; NID:g58405; PIDN:CAA43853.1; PID:g4494091

A:Note: The authors translated the codon TAC for residue 43 as Thr

C:Genetics:

A:Introns: 37/1

C:Superfamily: sperm histone

C:Keywords: chromosomal protein; DNA binding; nucleus; spermatogenesis

Query Match 2.5%; Score 7; DB 2; Length 50;

Best Local Similarity 100.0%; Pred. No. 28;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 ORRGRR 48

DB 18 ORRGRR 24

RESULT 22

S34045

protamine - North American opossum

C:Species: *Didelphis virginiana*, *Didelphis marsupialis virginiana* (North American OP

C>Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 23-Jul-1999

C:Accession: S34045

R:Winkfein, R.J.; Nishikawa, S.; Connor, W.; Dixon, G.H.

Eur. J. Biochem. 215, 63-72, 1993  
A:Title: Characterization of a marsupial sperm protamine gene and its transcripts from  
A:Reference number: S34045; MUID:93345500; PMID:8344286  
A:Accession: S34045  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-58 <MIN>  
A:Cross-references: EMBL:X74044; NID:9407062; PIDN:CAAS2193.1; PID:9407063  
C:Superfamily: sperm histone  
C:Keywords: DNA binding; nucleus

Query Match  
Best Local Similarity 100.0%; Score 7; DB 2; Length 58;  
Pred. No. 32;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 RRRGRG 49  
|||||  
Db 35 RRRGRG 41

RESULT 23  
F87604  
hypothetical protein CC2870 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: F87604  
R:Nieman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Lamb, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: F87604  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-86 <STO>  
A:Cross-references: GB:AE005673; NID:913424486; PIDN:AAK24834.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC2870

Query Match  
Best Local Similarity 100.0%; Score 7; DB 2; Length 86;  
Pred. No. 45;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 LALGLGL 65  
|||||  
Db 26 LALGLGL 32

RESULT 24  
F87993  
protein ZC334.3 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: F87993  
R:anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo  
A:Reference number: A75000; MUID:99068618; and www.sanger.ac.uk/projects/C\_ele  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_ele  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: F87993  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-102 <STO>  
A:Cross-references: GB:chr-I; PIDN:CAH04964.1; PID:93681432; GSPDB:GN00019; CESP:ZC334.3  
A:Note: predicted using GeneFinder  
C:Genetics:  
A:Gene: ZC334.3  
A:Map position: 1

Query Match  
Best Local Similarity 100.0%; Score 7; DB 2; Length 102;  
Pred. No. 52;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 LVLPLAL 61  
|||||  
Db 33 LVLPLAL 39

RESULT 25  
PH1560  
Ig heavy chain V region (clone VH32) - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 05-Aug-1994 #sequence\_revision 05-Aug-1994 #text\_change 21-Jan-2000  
C:Accession: PH1560  
R:Rassenti, L.Z.; Kipps, T.J.  
J. Exp. Med. 177, 1039-1046, 1993  
A:Title: Lack of extensive mutations in the VH5 genes used in common B cell chronic 1  
A:Reference number: PH1557; MUID:93210459; PMID:7681468  
A:Accession: PH1560  
A:Molecule type: DNA  
A:Residues: 1-115 <RAS>  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:32-115/Domain: immunoglobulin homology <IMM>

Query Match  
Best Local Similarity 100.0%; Score 7; DB 2; Length 115;  
Pred. No. 57;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 70 LGLLAV 76  
|||||  
Db 5 LGLLAV 11

RESULT 26  
H83201  
conserved hypothetical protein PA357 [imported] - Pseudomonas aeruginosa (strain PAC  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: H83201  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidis, K.; I  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa  
A:Reference number: A82950; MUID:20437337; PMID:10984043  
A:Accession: H83201  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-115 <STO>  
A:Cross-references: GB:AE004776; GB:AE004091; NID:99949701; PIDN:AA606945.1; GSPDB:GT  
A:Experimental source: strain PAO1  
C:Genetics:  
A:Gene: PA357

Query Match  
Best Local Similarity 100.0%; Score 7; DB 2; Length 115;  
Pred. No. 57;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 LALGLGL 65  
|||||  
Db 49 LALGLGL 55

RESULT 27  
IEEC5B  
hypothetical protein, 12K - Escherichia coli insertion sequence ISS  
C:Species: Escherichia coli  
C>Date: 18-Dec-1981 #sequence\_revision 18-Dec-1981 #text\_change 10-Sep-1999  
C:Accession: B91483; C03582; A04466  
R:Schoner, B.; Kahn, M.  
Gene 14, 165-174, 1981  
A:Title: The nucleotide sequence of ISS from Escherichia coli.  
A:Reference number: A91483; MUID:82028653; PMID:6269959

A:Accession: B91483  
 A:Molecule type: DNA  
 A:Residues: 1-118 <SCH>  
 R:Engler, J.A.; van Bree, M.P.  
 Gene 14, 155-163, 1981  
 A:Title: The nucleotide sequence and protein-coding capability of the transposable element  
 A:Reference number: A03582; MUID:82028652; PMID:6269958  
 A:Accession: C03582  
 A:Molecule type: DNA  
 A:Residues: 1-118 <ENG>  
 C:Genetics:  
 A:Mobile element: insertion sequence IS5  
 C:Superfamily: Escherichia coli hypothetical protein o263

Query Match 2.5%; Score 7; DB 1; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 DGGAVRQ 29  
 |||||  
 Db 107 DGGAVRQ 113

RESULT 28  
 AE1753  
 Orf51 [bacteriophage b1285] homolog 1ln2570 [imported] - Listeria innocua (strain Clp1  
 C:Species: Listeria innocua  
 C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001  
 A:Accession: AE1753  
 R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker  
 .; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fshl, H.  
 D.; Jones, L.M.; Karst, U.  
 Science 294, 849-852, 2001  
 A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maltournam, A.; Me  
 Ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlant,  
 A:Title: Comparative genomics of Listeria species.  
 A:Reference number: AB1077; MUID:21537279; PMID:11679669  
 A:Accession: AE1753  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-118 <GLA>  
 A:Cross-references: GB:AL592022; PIDN:GAC97797.1; PID:q16415092; GSPDB:GN00178  
 A:Experimental source: strain Clp11262  
 C:Genetics:  
 A:Gene: 1ln2570

Query Match 2.5%; Score 7; DB 2; Length 118;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SLIDFEI 8  
 |||||  
 Db 61 SLIDFEI 67

RESULT 29  
 AH2707  
 conserved hypothetical protein Atu1065 [imported] - Agrobacterium tumefaciens (strain C5  
 C:Species: Agrobacterium tumefaciens  
 C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
 A:Accession: AH2707  
 R:Wood, D.W.; Setubal, J.C.; Kaut, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L  
 erage, G.; Gillet, M.; Grant, C.; Guenther, D.; Kuyavlin, T.; Levy, R.; Li, M.; McCell  
 ; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A:Authors: Yoo, H.; Teo, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
 ster, E.W.  
 A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
 A:Reference number: AB2577; PMID:11743193  
 A:Accession: AH2707  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-123 <KOR>

A:Cross-references: GB:AE008688; PIDN:AA142078.1; PID:q17739458; GSPDB:GN00186  
 A:Experimental source: strain C58 (Dupont)  
 C:Genetics:  
 A:Gene: Atu1065  
 A:Map position: circular chromosome

Query Match 2.5%; Score 7; DB 2; Length 123;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 RLPLPRS 18  
 |||||  
 Db 50 RLPLPRS 56

RESULT 30  
 T27519  
 hypothetical protein ZC334.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 A:Accession: T27519  
 R:McLay, K.  
 submitted to the EMBL Data Library, November 1996  
 A:Reference number: Z20381  
 A:Accession: T27519  
 A:Status: preliminary; translated from GB/EMBL/DBD  
 A:Molecule type: DNA  
 A:Residues: 1-125 <WIL>  
 A:Cross-references: EMBL:Z82082; PIDN:CAB04964.2; GSPDB:GN00019; CESP:ZC334.3  
 A:Experimental source: clone ZC334  
 C:Genetics:  
 A:Gene: CESP:ZC334.3  
 A:Map position: 1  
 A:introns: 99/1

Query Match 2.5%; Score 7; DB 2; Length 125;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 LNVPLAL 61  
 |||||  
 Db 33 LNVPLAL 39

RESULT 31  
 I52290  
 interleukin-13 - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 16-Jul-1999  
 A:Accession: I52290  
 R:Liakis, F.G.; Gruet, E.N.  
 Biochem. Biophys. Res. Commun. 197, 612-618, 1993  
 A:Title: Cloning of rat interleukin-13 (IL-13) cDNA and analysis of IL-13 gene expre  
 A:Reference number: I52290; MUID:94092138; PMID:7916615  
 A:Accession: I52290  
 A:Status: preliminary; translated from GB/EMBL/DBD  
 A:Molecule type: mRNA  
 A:Residues: 1-131 <RES>  
 A:Cross-references: GB:I26913; NID:9438875; PIDN:AA16478.1; PID:9438876  
 C:Genetics:  
 A:Gene: IL-13  
 C:Superfamily: Interleukin-13

Query Match 2.5%; Score 7; DB 2; Length 131;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACG 71  
 |||||  
 Db 9 LALACG 15

RESULT 32

E30552  
T-cell activation protein P600 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 28-Aug-1989 #sequence\_revision 28-Aug-1989 #text\_change 16-Jul-1999  
C:Accession: E30552  
R:Brown, K.D.; Zurawski, S.M.; Mosmann, T.R.; Zurawski, G.  
J. Immunol. 142, 679-687, 1989  
A:Title: A family of small inducible proteins secreted by leukocytes are members of a new class of various activation processes.  
A:Reference number: A30552; MUID:89093958; PMID:2521353  
A:Accession: E30552  
A:Molecule type: mRNA  
A:Residues: 1-131 <BRO>  
A:Cross-references: GB:M23504; NID:9533246; PIDN:AAA40149.1; PID:9533247  
C:Superfamily: Interleukin-13

Query Match 2.5%; Score 7; DB 2; Length 131;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACIG 71  
|||||||  
DB 9 LALACIG 15

RESULT 33  
T37116  
Probable transposase, truncated [imported] - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C>Date: 08-Sep-2000 #sequence\_revision 08-Sep-2000 #text\_change 15-Sep-2000  
C:Accession: T37116  
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M. submitted to the EMBL Data Library, August 1999  
A:Reference number: 221588  
A:Accession: T37116  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-146 <SAU>  
A:Cross-references: EMBL:AL109950; PIDN:GAB52967.1; GSPDB:GN00070; SCOEDB:SCJ4.33c  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SCJ4.33c  
C:Superfamily: Synchocystis transposase s111710

Query Match 2.5%; Score 7; DB 2; Length 146;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLIAY 76  
|||||||  
DB 102 LGLLIAY 108

RESULT 34  
A71217  
Hypothetical protein PH2001 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
C:Accession: A71217  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Sekir  
M.; Ohtuku, Y.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
A:Reference number: A71000; MUID:98344137; PMID:9679194  
A:Accession: A71217  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-147 <KAW>  
A:Cross-references: GB:AP000007; GB:AP000001; NID:93236134; NID:93236128; PID:93258445;  
A:Experimental source: strain OT3  
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
C:Genetics:  
A:Note: this sequence is split into two separate translations in GenBank (residues 1-124

A:Gene: PH2001  
C:Superfamily: Pyrococcus horikoshii hypothetical protein PH2001

Query Match 2.5%; Score 7; DB 2; Length 147;  
Best Local Similarity 100.0%; Pred. No. 71;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 LACIGLL 73  
|||||||  
DB 45 LACIGLL 51

RESULT 35  
T08734  
Hypothetical protein DKFZp566F0546.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 13-Aug-1999  
C:Accession: T08734  
R:Ottenwelder, B.; Obermayer, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S. submitted to the Protein Sequence Database, May 1999  
A:Reference number: 216474  
A:Accession: T08734  
A:Molecule type: mRNA  
A:Residues: 1-150 <OTT>  
A:Cross-references: EMBL:AL050075  
A:Experimental source: fetal kidney; clone DKFZp566F0546  
C:Genetics:  
A:Note: DKFZp566F0546.1

Query Match 2.5%; Score 7; DB 2; Length 150;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 234 AASSLGP 240  
|||||||  
DB 32 AASSLGP 38

RESULT 36  
S31078  
Seed allergen RA5 - rice  
C:Species: Oryza sativa (rice)  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 20-Jun-2000  
C:Accession: S31078  
R:Adachi, T.; Izumi, H.; Yamada, T.; Tanaka, K.; Takeuchi, S.; Nakamura, R.; Matsuda, Plant Mol. Biol. 21, 239-248, 1993  
A:Title: Gene structure and expression of rice seed allergenic proteins belonging to  
A:Reference number: S31078; MUID:93144699; PMID:7678765  
A:Accession: S31078  
A:Molecule type: mRNA  
A:Residues: 1-157 <ADA>  
A:Cross-references: EMBL:D11430; NID:9218196; PIDN:BAA01966.1; PID:9218197  
C:Superfamily: wheat alpha-amylase inhibitor  
C:Keywords: seed

Query Match 2.5%; Score 7; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVVS 78  
|||||||  
DB 11 LLLAVVS 17

RESULT 37  
T02664  
Allergen - rice  
C:Species: Oryza sativa (rice)  
C>Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 16-Jul-1999  
C:Accession: T02664  
R:Yun, C.H.; Park, J.H.; Eun, M.Y. submitted to the EMBL Data Library, January 1998  
A:Description: Nucleotide sequence of rice allergenic protein.

A:Reference number: 214691  
A:Accession: F02664  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-157 <YUN>  
A:Cross-references: EMBL:AF042200; NID:g2827315; PIDN:AA099797.1; PID:g2827316  
A:Experimental source: strain N1pponbare  
C:Superfamily: wheat alpha-amylase inhibitor

Query Match 2.5%; Score 7; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVVS 78  
|||||||  
DB 11 LLLAVVS 17

RESULT 38  
A75567  
conserved hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000

C:Accession: A75567  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
S.; Shen, M.; Yamathavan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: A75567  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-157 <WHI>  
A:Cross-references: GB:AE001867; GB:AE000513; NID:g6457693; PIDN:AA09626.1; PID:g645769  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0033  
A:Map position: 1

Query Match 2.5%; Score 7; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 75;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 261 IRTLPWA 267  
|||||||  
DB 22 IRTLPWA 28

RESULT 39  
E75530  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000

C:Accession: E75530  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
S.; Shen, M.; Yamathavan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: E75530  
A:Status: preliminary  
A:Molecule type: DNA

A:Residues: 1-157 <WHI>  
A:Cross-references: GB:AE001865; GB:AE000513; NID:g6458024; PIDN:AA09935.1; PID:g645803  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR0352  
A:Map position: 1

Query Match 2.5%; Score 7; DB 2; Length 157;  
Best Local Similarity 100.0%; Pred. No. 75;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 43 RRRGRG 49  
|||||||  
DB 144 RRRGRG 150

RESULT 40

S59925  
allergen RA5B precursor - rice  
C:Species: Oryza sativa (rice)  
C:Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Jun-2000  
C:Accession: S59925

R:Alvarez, A.M.; Adachi, T.; Nakase, M.; Aoki, N.; Nakamura, R.; Matsuda, T.  
Biochim. Biophys. Acta 1251, 201-204, 1995  
A:Title: Classification of rice allergenic protein cDNAs belonging to the alpha-amyl  
A:Reference number: S59922; MUID:95399441; PMID:7669811  
A:Accession: S59925  
A:Status: preliminary  
A:Molecule type: mRNA

A:Residues: 1-160 <AIY>  
A:Cross-references: EMBL:D42142; NID:g1398917; PIDN:BA07713.1; PID:g1398918  
C:Superfamily: wheat alpha-amylase inhibitor

Query Match 2.5%; Score 7; DB 2; Length 160;  
Best Local Similarity 100.0%; Pred. No. 76;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVVS 78  
|||||||  
DB 11 LLLAVVS 17

RESULT 41

DNEC17  
outer membrane protein hlpA precursor - Escherichia coli (strain K-12)

N:Alternate names: DNA-binding 17K protein; histone-like protein hlp  
C:Species: Escherichia coli  
C:Date: 30-Jun-1989 #sequence\_revision 30-Jun-1989 #text\_change 01-Mar-2002  
C:Accession: J03034; A38063; S13728; B64742; I54944; S20426  
R:Holck, A.; Klepe, K.  
Gene 67, 117-124, 1988

A:Title: Cloning and sequencing of the gene for the DNA-binding 17K protein of Esche  
A:Reference number: J03034; MUID:88329735; PMID:2843433  
A:Accession: J03034

A:Molecule type: DNA  
A:Residues: 1-161 <HOI>  
A:Cross-references: GB:M21118; NID:g147821; PIDN:AA24630.1; PID:g147822  
A:Accession: A38063

A:Molecule type: protein  
A:Residues: 21-30 <HO2>  
A:Experimental source: strain B  
R:Dicker, I.B.; Seetharam, S.  
J. Bacteriol. 173, 334-344, 1991  
A:Title: Cloning and nucleotide sequence of the fira gene and the fira200(Ts) allele  
A:Reference number: S13728; MUID:91100302; PMID:1987124  
A:Accession: S13728

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 72-161 <DIC>  
A:Cross-references: EMBL:X54797; NID:g41468; PIDN:CAA38567.1; PID:g41469

A:Experimental source: strain K-12, substrain MG1655  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: B64742

A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA

A:Residues: 1-161 <BLAT>  
A:Cross-references: GB:AE000127; GB:U00096; NID:g1786370; PIDN:AA073289.1; PID:g1786  
A:Experimental source: strain K-12, substrain MG1655

R.Hirvas, L.; Koski, P.; Vaara, M.  
 J. Bacteriol. 173, 1223-1229, 1991  
 A:Title: The ompH gene of *Yersinia enterocolitica*: cloning, sequencing, expression, and  
 A:Reference number: 154944; MID:91123198; PMID:1991717  
 A:Accession: 154944  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-14, 'L', 16-148, 'E', 150-152, 'I', 154-161 <RES>  
 A:Cross-references: EMBL:X75465; NID:9432661; PIDN:CA53207.1; PID:9432662  
 C:Comment: The hlpA protein has been believed to be a histone-like constituent of bacteria  
 C:Genetics:  
 A:Gene: hlpA, sfp  
 C:Superfamily: DNA-binding 17K protein  
 C:Keywords: membrane protein  
 F:1-20/Domain: signal sequence  
 F:21-161/Product: outer membrane protein hlpA #status predicted <STG>

Query Match 2.5%; Score 7; DB 1; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 GIGLALA 68  
 |||||  
 DB 9 GIGLALA 15

RESULT 42  
 D90651  
 histone-like protein HlpA [imported] - *Escherichia coli* (strain O157:H7, substrain RMD  
 C:Species: *Escherichia coli*  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
 A:Accession: D90651  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
 gerasavara, N.; Yasunaga, T.; Kuhara, S.; Shibata, T.; Hattori, M.; Shingawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and gene  
 A:Reference number: A95629; MID:21156231; PMID:11258796  
 A:Accession: D90651  
 A:Molecule type: DNA  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-161 <HAY>  
 A:Cross-references: GB:BA000007; PIDN:BA833603.1; PID:913359636; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RMD 0509952  
 C:Genetics:  
 A:Gene: EC80180  
 C:Superfamily: DNA-binding 17K protein

Query Match 2.5%; Score 7; DB 2; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 GIGLALA 68  
 |||||  
 DB 9 GIGLALA 15

RESULT 43  
 D85502  
 hypothetical protein hlpA [imported] - *Escherichia coli* (strain O157:H7, substrain EDL93  
 C:Species: *Escherichia coli*  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 A:Accession: D85502  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Diallanita, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
 A:Reference number: A85480; MID:21074933; PMID:11206551  
 A:Accession: D85502  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-161 <SNO>  
 A:Cross-references: GB:AE005174; NID:912512906; PIDN:AAG54480.1; GSPDB:GN00145; UWGP:Z01  
 A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:  
 A:Gene: hlpA  
 C:Superfamily: DNA-binding 17K protein

Query Match 2.5%; Score 7; DB 2; Length 161;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 GIGLALA 68  
 |||||  
 DB 9 GIGLALA 15

RESULT 44  
 T24937  
 hypothetical protein W03C9.4 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 29-Oct-1999  
 A:Accession: T24937; T26123  
 R:Gregory, J.  
 submitted to the EMBL Data Library, July 1995  
 A:Reference number: T24937  
 A:Accession: T24937  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-162 <WIL>  
 A:Cross-references: EMBL:Z50015; PIDN:CA90314.1; GSPDB:GN00020; CESP:W03C9.4  
 A:Experimental source: clone T15G9  
 R:Gregory, J.; Alnscough, R.  
 submitted to the EMBL Data Library, October 1995  
 A:Reference number: Z20155  
 A:Accession: T26123  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-162 <WIL>  
 A:Cross-references: EMBL:Z66516; PIDN:CA91361.1; GSPDB:GN00020; CESP:W03C9.4  
 A:Experimental source: clone W03C9  
 C:Genetics:  
 A:Gene: CESP:W03C9.4  
 A:Map position: 2  
 A:Introns: 16/3; 64/1; 120/3; 128/1

Query Match 2.5%; Score 7; DB 2; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 77;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 231 SATASS 237  
 |||||  
 DB 112 SATASS 118

RESULT 45  
 T31173  
 hypothetical protein 424 - *Sphingomonas aromaticivorans* plasmid pNLI  
 C:Species: *Sphingomonas aromaticivorans*  
 C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jun-2000  
 A:Accession: T31173  
 R:Romine, M.F.; Stillwell, L.C.; Wong, K.K.; Thurston, S.J.; Sisk, E.C.; Sensen, C.W.  
 submitted to the EMBL Data Library, July 1998  
 A:Description: Complete sequence of a 184 kb catabolic plasmid from *Sphingomonas* *arom*  
 A:Reference number: Z20992  
 A:Accession: T31173  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-162 <ROM>  
 A:Cross-references: EMBL:AF079317; NID:93378261; PID:93378314; PIDN:ADD03897.1  
 C:Genetics:  
 A:Genome: plasmid pNLI  
 A:Note: ori424  
 C:Superfamily: *Sphingomonas aromaticivorans* hypothetical protein 424  
 Query Match 2.5%; Score 7; DB 2; Length 162;  
 Best Local Similarity 100.0%; Pred. No. 77;

	Matches	7;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Oy	138	ARRATA	144							
Db	28	ARRATA	34							

Search completed: May 8, 2003, 07:07:55  
Job time : 50 secs

---

**THIS PAGE BLANK (USPTO)**



From: Chan, Christina  
Sent: Monday, May 05, 2003 3:41 PM  
To: Schnizer, Richard; STIC-Biotech/ChemLib  
Subject: RE: 09/245,198

Please rush. Thanks Chris

Chris Chan  
TC 1600 New Hire Training Coordinator and SPE 1644  
308-3973  
CM-1, 9B19

-----Original Message-----

From: Schnizer, Richard  
Sent: Monday, May 05, 2003 3:21 PM  
To: Chan, Christina  
Subject: 09/245,198

Please authorize the following RUSH search for 09/245,198 which is an amended case in condition for allowance.

Please search the commercial and interference databases for:

-SEQ ID NOS:1-4, and

-nucleic acids that could encode the amino acid sequences of SEQ ID NOS:2 and 4

Please perform oligo searches on SEQ ID NOS: 3 and 4.

Thank you-

Richard Schnizer, Ph.D.  
Patent Examiner  
Art Unit 1635  
CM1 12E17  
703-306-5441  
Mail Box CM1 11E12

RECEIVED  
MAY -5 2003  
STIC/CHEM DIVISION

Point of Contact:  
Thomas G. Larson, Ph.D.  
703-308-7309  
CM1, Rm. 6 B 01

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 5/6  
Date Completed: 5/12  
Searcher Prep/Review: 10  
Clerical: 10  
Online time: 10

TYPE OF SEARCH:  
NA Sequences: 3  
AA Sequences: 2  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**